The Dissertation Committee for Rachel Michelle James Certifies that this is the approved version of the following dissertation:

THE DEVELOPMENT OF A HOLISTIC APPROACH TO MODELING DRIVER BEHAVIOR: ACCOUNTING FOR DRIVER HETEROGENEITY IN CAR-FOLLOWING MODELS

Committee:

Stephen Boyles, Supervisor

Randy Machemehl

S. Travis Waller

Zhanmin Zhang

THE DEVELOPMENT OF A HOLISTIC APPROACH TO MODELING DRIVER BEHAVIOR: ACCOUNTING FOR DRIVER HETEROGENEITY IN CAR-FOLLOWING MODELS

by

Rachel Michelle James

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2019

Dedication

For my grandparents, Joe and Robin Hale. Thank you for being my rock. At a young age, you instilled in me that education was my one-way ticket to a better future. I know you made sacrifices to give me a future that was unavailable for you and I appreciate that immensely. I love you the most, forever and always.

Acknowledgements

First and foremost, I want to thank my dissertation committee: Dr. Boyles, Dr. Machemehl, Dr. Zhang, and Dr. Waller. I sincerely appreciate your time, effort, and support throughout this journey. In particular, to my research supervisor, Dr. Boyles: I appreciate your patience and your willingness to let me try something different for my doctoral journey. I appreciate you taking a chance on me five years ago and hope that we will find ways to collaborate in the future.

To all of the educators who have encouraged my love of learning: thank you. Education truly is a channel for upward mobility and I appreciate the role you serve in changing lives, including mine. In particular, to Dr. Avi: you deserve so much credit for putting me on the PhD path when you invited me to join your lab as an undergraduate researcher. I will forever be grateful for all of the opportunities you provided me as your student and for being there for the occasional pep talk when I hit a massive roadblock.

To my group fitness community in Austin and Arlington: Thank you for being such a great distraction throughout this educational marathon! In particular, to my Saturday Coffee Club—Megan, Greg, Katie, Karin, Jess, Melissa, Christine, Chrissie, Andrew, Mike, Katherine, and Bob, among others—I love you. Thank you for being so wonderful.

To all of my MS and PhD friends at UT: Thank you for being such a wonderful source of comedic relief. You all are so incredibly talented, intelligent, and will be powerful forces of change. To Kristie, John, and Venktesh: you three are so special to me. Our time together will be what I miss most about graduate school.

To the staff at the Turner-Fairbank Highway Research Center: Turner-Fairbank is such a special place. Specifically, I want to acknowledge the Office of Operations Research and Development. You have been so invested in my success since day one and I appreciate

all of the opportunities you have provided. To the leadership team—Brian, Gene, and Dale—I am eternally grateful to you for giving me the flexibility and freedom to grow as a professional while completing my dissertation. To all of the visiting researchers—especially Britton, Yi, Mehdi, Joe, Josh, and Pavle—thank you for being willing to review papers, brainstorm in the Innovation Lab, the CAVA (x3) runs, and game nights.

To my unofficially adopted family, the Vestris–Greg, Jean, Rachel, Ariel, Drew, Jocelyn, Haley, Addie, Tyler, Jacob, and Wallace: Thank you for loving me so well. I appreciate the unconditional love and your total investment in my happiness and success. You are all so dear to me.

To the Murray family–Roger, Mary Ann, Sarah, Brink, Lenn, Brittany, Elliott, Westyn, Cora, and August: I cannot properly thank you for being so supportive. Thank you for accepting me into your family with wide-open arms. I love you all so much.

To my siblings–Adam, Brian, Michael, Isaac, and Elijah: One of my biggest goals in life is to be the role model you deserve. I hope this shows you that your past does not dictate your future and that you can achieve anything with hard work and a bit of grit.

To Britton Elaine: I could almost double the length of this document with all of things I appreciate about you. I will try and distill it into a few words: thank you for giving me the courage to travel; thank you for the many coding homework assignments and Python lessons along the way; thank you for constantly being willing to help to debug and finalize code; thank you for being the best research partner in the universe; and thank you for being such a wonderful source of pick-me-ups and pep talks. I have gained so much from this doctoral journey: your friendship and sisterhood are the greatest gift of all.

Finally, to Ian and my fur babies, Minnie and Lily: Your love and support have been foundational to the completion of this degree. I love you so much. This would not have been possible without you.

The Development of a Holistic Approach to Modeling Driver Behavior: Accounting for Driver Heterogeneity in Car-Following Models

Rachel Michelle James, Ph.D.

The University of Texas at Austin, 2019

Supervisor: Stephen Boyles

Car-following behavior has been studied since the 1940s. However, complex calibration requirements and challenges with collecting high-resolution data have stunted advancements in this domain. Thus, methodologies to adequately capture naturalistic behavioral heterogeneity are largely missing from the literature.

For this dissertation, a sample from the second Strategic Highway Research Program Naturalistic Driving Study was analyzed. This sample contains 665 trips completed on freeways in clear weather conditions. Driver demographics, vehicle CAN bus, and external sensor data are available for each trip. The trajectories in this sample were processed and used to calibrate the Gipps, Intelligent Driver Model, and Wiedemann 99 car-following models.

This dissertation seeks to improve how inter-driver heterogeneity in car-following behavior is accounted for in microsimulation models. This dissertation has three primary objectives. Objective 1 identifies which driver attributes are sources of inter-driver heterogeneity. Objective 2 explores the viability of using census-level data to calibrate microsimulation models. Objective 3 develops and evaluates a new mechanism for properly capturing inter-driver heterogeneity in microsimulation: an ensemble car-following model.

To achieve these objectives, first, Kruskal-Wallis one-way analysis of variance tests were applied to show statistically significant differences in both the estimated car-following model calibration coefficients and the overall model performance across groups of drivers categorized by commonalities in their driver attributes.

Next, the *Expectation Maximization* clustering algorithm was applied to show that, despite differences in driver behavior, homogeneous driver groups, or groups of drivers that behave similarly, exist in the dataset. Moreover, this dissertation shows that drivers can be classified into their proper homogeneous driver group only knowing their driver specific attributes.

Finally, VISSIM was used to implement the homogeneous driver groups in microsimulation. This case study illustrated that when inter-driver differences in driving behavior are explicitly modeled, there are notable impacts on the performance metrics collected from the microsimulation models. These performance metrics are ultimately used by decision makers to evaluate alternatives for transportation funding. Thus, this dissertation provides evidence of the importance of appropriately modeling inter-driver differences to improve the quality of the microsimulation model results and inform better funding allocation decisions.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Individual preferences in driving behavior (e.g., speed, reaction time) are highly variable between and within drivers; in simpler terms, different drivers drive differently. In fact, it is common for groups of people to argue over whom is the best driver because, despite the statistical impossibility, the majority of drivers believe they are an above average driver (Roy & Liersch, 2013). Yet, the field of microsimulation modeling does little towards systematically modeling these known differences in driving behavior. This dissertation seeks to fill this void by exploring the heterogeneity in driving behavior attributable to driver specific attributes, such as age and gender.

First, this dissertation introduces the idea of "homogeneous driver groups." Homogeneous driver groups are developed through the application of statistical tests, clustering algorithms, and classification algorithms to identify groups of drivers that exhibit behavior that is sufficiently similar within a group of like drivers, but sufficiently different from behavior observed in other groups. Next, this dissertation calibrates car-following models to match the behavior observed in the different homogeneous driver groups using data collected through the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS). This dissertation then explores how different proportions of homogeneous driver groups in the traffic stream impact the outputs of microsimulation models and explores the concept of using census-level data to calibrate the behavioral component of microsimulation models; this effectively moves the behavioral calibration to the back-end of the effort and significantly increases the practicality of accounting for behavioral heterogeneity in car-following models. Finally, this dissertation develops the idea of an "ensemble" car-following model, which uses

multiple diverse car-following models together to better capture the inherent variability in naturalistic data.

1.1. MOTIVATION

Transportation planning is the comprehensive analysis of developing the vision and goals for a region, quantitatively evaluating and prioritizing projects/strategies, monitoring the implementation and operation of the system, and using the data and lessons learned to make better future investment decisions (Federal Highway Administration Office of Planning, 2007). It is a cooperative process between the local Metropolitan Planning Organization (MPO), state Department of Transportation (DOT), transit operating agency (where applicable), and the public.

Traffic analysis tools are designed to transparently inform the decisions that feed the transportation planning process (Jeannotte, Chandra, Alexiadis, & Skabardonis, 2004). These analyses occur on a multitude of scales: microscopic, mesoscopic, and macroscopic. Microscopic models simulate the realistic movement of individual vehicles through the network; they are frequently composed of car-following (acceleration/deceleration), lane-changing, gap acceptance, speed selection, and other modules that represent the grander driving task (Alexiadis, Jeannotte, & Chandra, 2004). These extremely detailed models enable the detailed analysis of the tradeoffs between projects, but are limited in spatial scope, challenging to calibrate, and are computationally expensive to run. Macroscopic models are based on the well-established relationships between traffic flow, density, and speed; the reduction in details enables the expansion of model scope with the tradeoff of model accuracy. Mesoscopic models exhibit properties of both microscopic and macroscopic models. In mesoscopic models, the relationships between traffic density,

volume, and flow are respected, as with macroscopic models; however, traffic is modeled at the link or "cell" level, which offers increased accuracy and resolution of the results, though not to the degree of microscopic models. This dissertation is primary focused on improving the methods and tools used in microscopic traffic analyses. Different communities within the transportation profession use these models in different ways. It is anticipated that this dissertation will be of most value to the traffic flow theory community of researchers, as this dissertation explores the impact of modeled heterogeneity on characteristics of traffic flow (e.g., jam density, speed at capacity, capacity).

With the enactment of Moving Ahead for Progress in the 21st Century (MAP-21), there has been an increase in emphasis on performance-based planning and scenario development for federally funded transportation investments (Federal Highway Administration, 2013). Moreover, the most recent Surface Transportation Reauthorization Act, the Fixing America's Surface Transportation (FAST) Act, placed increase emphasis on the application of traffic analysis tools in the planning process (114th Congress, 2015). Section 1430 explicitly states:

> "The Department should utilize, to the fullest and most economically feasible extent practicable, **modeling and simulation technology to analyze highway and public transportation projects** authorized by this Act to ensure that these projects—"
>
> (1) will increase transportation capacity and safety, alleviate congestion, and reduce travel time and environmental impacts; and
>
> (2) are as cost effective as practicable." (Fixing America's Surface Transportation Act, 2015, p. 117)

Thus, the accuracy and robustness of modeling tools is of increased interest to transportation agencies, consultants, and researchers alike. Microsimulation software can

be used as a powerful tool in scenario analyses, as the outputs are an explicit representation of the different decisions of an individual driver. However, increased resolution of model outputs requires the modeler to ensure simulation accuracy in producing reasonable approximations of individual driver behavior. Outputs of these models are used to inform multimillion-dollar investments in transportation; thus, the realism of the inputs and algorithms housed within a model are of upmost importance.

One of the biggest limitations of microscopic simulation models is the immense difficulty in properly calibrating the model (Jeannotte et al., 2004). Calibration is the process by which the modeler adjusts pre-defined model parameters in order to simulate the traffic performance of a facility. Inputs to microsimulation models for calibration include geometric data, control data, demand data, and system performance data (Dowling, Skabardonis, & Alexiadis, 2004). Of particular interest to this dissertation is the required system performance data. Traditionally, these data has included travel times, delays, queues, speeds, and traffic counts aggregated across varying temporal resolutions. Thus, modelers are using macroscopic data to calibrate microscopic parameters (e.g., standstill acceleration and headway time in VISSIM; maximum desired acceleration and maximum desired deceleration in the Intelligent Driver Model (IDM); reaction time and desired velocity in Gipps, etc.).

This approach, though reflective of state-of-practice and found to be "good enough", is fundamentally flawed. It can be easily argued that microscopic data points can be aggregated into macroscopic summary values and tell a consistent story; the inverse is not necessarily true. Take a set of five numbers as an example (e.g., [10,10,10,10,10]). These numbers can be aggregated to tell a bigger picture (e.g., the average of this particular set of numbers is 10 and can only ever be 10). However, if it is known that the average of five numbers is 10, there are numerous combinations of five numbers that can be used to

achieve a mean of 10. This is essentially what we, as a practice, are supporting when calibration best practices include iteratively changing coefficients representing microscopic parameters to match macroscopic data.

However, recent advances in data collection and storage capabilities—via instrumented vehicles and aerial videos—have enabled the collection of calibration data that actually match the scope of the microscopic model: the vehicle trajectories themselves. Programs like the SHRP2 NDS (Virginia Tech Transportation Institute, 2018b) and more recent projects sponsored by the Federal Highway Administration Office of Research, Development, and Technology (Federal Highway Administration Office of Operations Research and Development, 2017b) have created the conditions for an overhaul of the microscopic simulation model calibration state-of-practice in favor of a procedure that is more transparent in nature and uses naturalistic position, velocity, and acceleration data to calibrate driver behavior parameters.

Moreover, the more accurate representation of driver behavior—such as instantaneous acceleration, desired gap, and relative velocity—enable researchers to advance newer fields such as emissions modeling (Chen & Yu, 2007; Madi, 2016; Song, Yu, & Zhang, 2012; Stevanovic, Stevanovic, Zhang, & Batterman, 2009) and the modeling of the interaction of human driven vehicles with automated driving systems (ADS) (James, Melson, Hu, & Bared, 2018; Ma, Zhou, Huang, & James, 2018). Though common microsimulation software packages have started integrating with simulation modules for emissions and ADS, multiple research efforts have indicated that the (human) driver behavior in the microsimulation software is not consistent with naturalistic driving data; this significantly limits the realism of research efforts that rely on these programs to accurately characterize vehicle dynamics as a function of human factors (Andrew L Berthaume, 2015; Yang & Morton, 2012).

A secondary avenue for research with more accurate driver model capabilities are risk-based crash analyses (Fan, Wang, Liu, & Yu, 2013; Habtemichael & De Picado Santos, 2014). In 2008, the Federal Highway Administration sponsored the development of the Surrogate Safety Assessment Model (SSAM) software (Gettman, Pu, Sayed, & Shelby, 2008). This software was designed to enable risk-based crash analysis via the extraction of vehicle trajectories from microsimulation software such as VISSIM, CORSIM, and ETFOMM. However, a vast majority of microsimulation software packages are designed to elicit unrealistic acceleration values to avoid simulation collisions (Xyntarakis, Alexiadis, Campbell, & Flanigan, 2016); this completely circumvents the use of microsimulation vehicle trajectories as a tool to understand safety. Through the development and calibration of car-following models based on naturalistic data, it is much more likely that meaningful analyses may be completed to understand the impact of operational and geometric changes on the overall safety level of a facility.

Though the field of car-following has been studied since the 1940s, several limitations of data collection, complications with calibration, and lack of model transparency have severely stunted the development of accurate methods in this domain. This dissertation seeks to contribute to the field via the development of data-driven methods to better account for driver heterogeneity in the car-following process.

1.2. DISSERTATION CONTRIBUTIONS

This dissertation is a broad exploration of heterogeneity in driving behavior captured in trajectory-level data. Moreover, this dissertation investigates the degree to which heterogeneity in driving behavior should be captured in transportation planning tools, specifically car-following models. As such, this dissertation uses trip statistics and

calibrated parameter sets as surrogates for driver behavior. In order to process the data to obtain sets of estimated calibrated parameter coefficients, three steps were required: (i) the identification of constrained driving (i.e., car-following) states from time-series processed radar data using a radar-vision algorithm, discussed in Section 3.2.2; (ii) the identification of candidate car-following models, covered in Section 3.2.3; and (iii) the identification of optimal car-following model parameter sets, discussed in Section 3.2.4.

The primary contributions of this dissertation are not the aforementioned data processing tasks, but rather analyses conducted with the processed data. The work conducted as part of this dissertation is split into five tasks and eight research questions:

- Task 1: Evaluate the presence of heterogeneity in the SHRP2 NDS dataset and determine if existing car-following models in the literature are suitable for capturing this heterogeneity at a trip-level.
    - Research Question 1: Is there evidence of driving behavior heterogeneity in the SHRP2 NDS time-series data?
    - Research Question 2: Can existing car-following models appropriately capture the diverse driving behavior of different drivers in the sample of SHRP2 NDS?
    - Research Question 3: Which model(s) best capture the diverse driving behavior recorded in trajectory-level data?
- Task 2: Characterize the inter-driver heterogeneity evident in the SHRP2 NDS as a function of driver specific attributes (e.g., age, gender).
    - Research Question 4: Do different hypothesized groups of drivers exhibit statistically significant differences in driving behavior?

    - Research Question 5: Do different subgroups of drivers behave sufficiently similarly to be considered one homogeneous group of drivers (i.e., do homogeneous driver groups exist in trajectory-level data?)
- Task 3: Identify method(s) to capture the collective behavior of homogeneous driver groups in microsimulation models.
    - Research Question 6: What methods should be used to obtain a representative set of calibration coefficients for a group of drivers?
- Task 4: Evaluate the utility of accounting for driver attributes in microsimulation.
    - Research Question 7: Can census-level (i.e., driver demographics) data be used alongside the anticipated proportions of driver subgroups to calibrate the car-following behavior of microsimulation models, effectively moving the calibration process to the back-end?
- Task 5: Evaluate a new method for capturing inter-driver heterogeneity: an "empirical" driver model.
    - Research Question 8: Does the diverse driving behavior observed in trajectory-level data require the application of multiple car-following models to more adequately capture the apparent heterogeneity in driving styles?

1.3. ORGANIZATION OF DISSERTATION

Chapter 1 provides motivation for this dissertation effort, as well as a roadmap for how this dissertation is organized. Chapter 2 provides a substantial review of literature on the topics of microsimulation models and documented driving behavior heterogeneity. Chapter 3 provides a detailed summary of the data processing used to obtain the sets of

car-following model calibration coefficients for analysis in this dissertation. This includes a summary of the SHRP2 data collection effort, the Wyoming Implementation Assistance Program, the extraction of car-following states from time-series radar data, and the developed calibration procedure. Chapter 4 provides a high-level overview of the calibration results, intended to provide evidence of unexplained heterogeneity in the calibrated model parameter estimates. Chapter 5 explores the viability of using driver specific attributes to segment the data into subgroups with reduced behavioral heterogeneity. This chapter specifically explores the differences between segmented groups of drivers, similarities within segmented groups of drivers, and begins to cultivate the idea of homogeneous driver groups. Chapter 6 explores the calibrated car-following model performance (i.e., calibration scores) across different driver attributes. Chapter 7 explores methods to obtain representative sets of calibration coefficients to describe the collective behavior of a group of drivers.

Chapter 8 and 9 put into practice the methodologies developed in Chapter 3, 4, 5, 6, and 7. Chapter 8 develops and applies a new framework for microsimulation model calibration, which relies on census-level data to calibrate the car-following model component of microsimulation models. Chapter 9 develops the first "ensemble" car-following model, which leverages multiple recognized car-following models in the literature to better characterize the diverse driving behavior evident in naturalistic driving data. Finally, concluding remarks, limitations of this dissertation research, and future research paths are discussed in Chapter 10.

Chapter 2: Literature Review

This chapter details topics fundamental to the proposed dissertation research area. Section 2.1 covers information related to the collection of microscopic trajectory-level data. Section 2.2 reviews popular car-following models in the literature. Section 2.3 covers literature pertaining to driver heterogeneity in empirical data (Section 2.3.1) and efforts to model heterogeneity in car-following behavior (Section 2.3.2). Section 2.4 briefly summarizes the key takeaways from Chapter 2.

2.1. MICROSCOPIC TRAJECTORY-LEVEL DATA COLLECTION METHODOLOGIES

Microscopic trajectory-level data is required to truly understand the human factors impacts on vehicle dynamics. Microscopic trajectory-level data involves the collection of detailed vehicle data—such as position, velocity, and acceleration—at a frequency of 1Hz (once a second) or higher (e.g., 10Hz, which is ten times a second). Microscopic data can be collected from numerous sources: instrumented research vehicle (IRV), driving simulator, and aerial video collection. Each of these methodologies have advantages and disadvantages for their use; however, the relationship between the different types of microscopic data are not well understood and no guidance currently exists to aid agencies in identifying which data collection method best suits their needs. This section details previous research efforts that collected microscopic trajectory-level data for research studies.

2.1.1. Instrumented Research Vehicle Data

An instrumented research vehicle, also referred to as instrumented personal vehicle in naturalistic studies, is an advanced method available to collect situational microscopic

trajectory-level data. This typically involves the inconspicuous instrumentation of a vehicle to collect forward- and rear-facing video; vehicle Controller Area Network bus (CAN bus) data; and forward-facing radar. This enables the collection of the equipped vehicle's position, velocity, and acceleration, as well as the relative position and velocity of adjacent vehicles (McLaughlin, Hankey, & Dingus, 2009). The primary benefit of long-term naturalistic driving studies collected via IRV is that the data collected represents real-world, routine driving; although radar data are inherently noisy, they appear to be less error ridden than data collected aerially, where any errors in the point measures of vehicles' location compound as velocity and acceleration estimates are derived (Punzo, Borzacchiello, & Ciuffo, 2011). Additionally, rich information about each driver is available and can be mapped to the behavioral data. The immense challenge of setting up the project and managing the overwhelming amount of data are disadvantages of IRV data collection efforts. Additionally, IRV forward-facing video and radar only provides details about the immediate downstream vehicle, limiting the analysis to one leading vehicle (i.e., IRV collected data is unable to study the topic of multi-attention in driving behavior). Examples of prior efforts to collect trajectory-level data with IRVs include the Bosch Research Group data collection; 2002 Naples, Italy IRV data collection; Virginia Tech Transportation Institute Naturalistic Truck and Car Studies; Federal Highway Administration Office of Research, Development, and Technology Living Laboratory; and the second Strategic Highway Research Program Naturalistic Driving Study.

Robert Bosch GmbH data collection effort

The Robert Bosch GmbH research group collected vehicle trajectory data during afternoon peak stop-and-go conditions on a single lane in Stuttgart, Germany in 1995 (Manstetten, Krautter, & Schwab, 1997). An IRV collected relative speed, relative

headway, and following vehicle acceleration. This data was recorded in 100ms intervals. Three of these trajectories, with car-following durations of 250s, 300s, and 400s, are available online and have enabled multiple efforts to apply trajectory-level data to calibrate car-following models (Kesting & Treiber, 2008; Panwai & Dia, 2005b, 2005a).

Naples, Italy data

Trajectory-level data was obtained from a series of experiments conducted in Naples, Italy between October 2002 and July 2003. The experiment consisted of four IRVs following in a platoon along urban and suburban roads under varying traffic conditions. The trajectory data was recorded at a frequency of 10Hz. This data enabled the study of both intra- and inter-driver heterogeneity (Papathanasopoulou & Antoniou, 2015; Papathanasopoulou, Markou, & Antoniou, 2016; Punzo, Ciuffo, & Montanino, 2012; Punzo & Simonelli, 2005)

Virginia Tech Transportation Institute Naturalistic Truck and Car Data Collection Studies

The Virginia Tech Transportation Institute (VTTI) sponsored three large naturalistic driving study data collection efforts for both heavy vehicles and personal vehicles (Abbas et al., 2012): (i) the Drowsy Driving Warning System Field Operational Test (DDSW FOT), (ii) the Naturalistic Truck Driving Study (NTDS), and (iii) the 100-Car Naturalistic Study. The data acquisition system included front radar, forward- and rearward-facing camera, and in-vehicle cameras. Position, velocity, and acceleration data was collected at a 10Hz frequency.

The VTTI DDWS FOT collected data from 103 truck drivers using 46 instrumented heavy vehicles. Each driver drove 60 hours a week for an average of 12 weeks. This

resulted in approximately 48,000 hours of driving data spanning 2.2 million miles (Abbas et al., 2012). This study produced 1,271 safety critical events.

The VTTI NTDS collected data for 100 commercial motor vehicle drivers for four months. Approximately 14,600 hours of driving data were collected traversing 735,000 miles (Abbas et al., 2012). More than 2,800 safety critical events were identified.

The VTTI 100-Car Naturalistic Study instrumented 100 light vehicles for naturalistic data collection. Continuous driving data were collected for 108 individual drivers; this resulted in nearly 43,000 hours of video across two million miles (Abbas et al., 2012). This was the first study to instrument personally owned vehicles to collect large-scale, naturalistic driving data.

These efforts collected trajectory-level data, with keen interest in safety critical events (e.g., to identify contributing factors of near crash events). However, later studies explored the viability of using this detailed trajectory-level data to explore situational driving behavior in baseline events (Abbas, Higgs, Adam, & Medina, 2011; Higgs & Abbas, 2015; Higgs, Abbas, & Medina, 2012)

Federal Highway Administration Living Laboratory

The Federal Highway Administration (FHWA) Office of Research, Development, and Technology (RD&T) sponsored an extensive data collection effort to study inter- and intra-driver heterogeneity between work zones and non-work zones through the use of a Living Laboratory in Northern Virginia (Federal Highway Administration Office of Operations Research and Development, 2017a). A living laboratory, defined in the context of transportation operations, is a transportation network instrumented with technology for the collection of user-centric data for evaluation of operational performance (Lochrane, Al-Deek, Dailey, & Bared, 2014). In this case, both an IRV and a connected mobile traffic

sensing (CMTS) system were developed for the collection of trajectory data on a freeway with both work zone and non-work zone segments.

A sport utility vehicle was equipped with the necessary on-board equipment: two universal median range radars (UMRR), for the collection of relative velocity and position of adjacent vehicles every 40ms; a speed sensor, to provide speed data at a frequency of 10Hz; a 30 frames per second video recording system; and a computer for the data acquisition. For additional details on equipment specs, see Lochrane, 2014.

The living laboratory to collect this data was located along the I–95 corridor outside of Washington, D.C. stretching from Arlington, VA to Springfield, VA; this included both work zone and non-work zone sections. A total of 64 participants drove approximately 50 miles—two to three hours—in either morning or afternoon peak. The participants were split equally in gender and ranged in age from 18 to 71, with a mean age of 41.4. These data were collected in 2013.

Because the data from the different equipment were collected at different frequencies, the data was consolidated to be reported every 0.1 second (10Hz). A filter was applied to remove noise from the data. A car-following period was specifically defined as an event in which the leading vehicle was less than 60 meters from the following vehicle in space and an event that lasted more than 20 seconds in time; later post-processing identified and removed hook-following. The final subset of data included approximately 900 miles and 1.65 hours of car-following events. These data were used to create a new multidimensional psychophysical framework for modeling driver behavior (Lochrane, Al-Deek, Dailey, & Krause, 2015); to inform the creation of a new social force car-following model, Modified Field Theory (Andrew L Berthaume, 2015); and to inform the creation of the FHWA Driver Model Platform, an open source graphical user interface designed to

streamline the use of data-driven car-following models in simulation software (Federal Highway Administration Office of Operations Research and Development, 2017a).

The second Strategic Highway Research Program Naturalistic Driving Study

The second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) is the largest study of its kind and was intended to collect data to better understand the role of driver behavior in traffic safety (Campbell, n.d.). Over 3,400 drivers participated in six different geographical locations. The participants were both male and female and ranged in age from 16 to 76+. This produced over 5.4 million trip summary records and more than 36,000 crash, near crash, and baseline driving events. Additional information on the SHRP2 NDS dataset is available through Insight (Virginia Tech Transportation Institute, 2018b). This dissertation specifically focuses on data collected via the SHRP2 NDS.

2.1.2. Aerial Data

Aerial data involves the extraction of vehicle trajectories from consecutive still images or videos. This has been accomplished using video cameras on tall buildings (Halkias & Colyar, 2006), via helicopter (Netherlands Organisation for Scientific Research, n.d.), and via drone (Federal Highway Administration Office of Operations Research and Development, 2017b). The benefit of aerial photography is that it allows the collection of multiple vehicle trajectories simultaneously with the prevailing traffic conditions. However, driver attributes are not identifiable, the data tends to be collected for short spatial intervals, and many data errors have been reported with aerial data.

Netherlands Organization for Scientific Research

One of the foundational efforts toward characterizing driver heterogeneity in car-following using microscopic data is the *Tracing Congestion Dynamics: With Innovative Microscopic Data to a Better Theory* sponsored by the Netherlands Organization for Scientific Research (TNO) (Netherlands Organisation for Scientific Research, n.d.). This project was awarded to Dr. S. P. Hoogendoorn at the Technische Universiteit Delft and spanned nine years, from February 2004 to May 2015. This project funded the collection of microscopic trajectory-level data via helicopter; the data was mined using remote sensing techniques. The data was recorded at a frequency of 10Hz over 400–500m of roadway for Freeway A2 in Urecht, the Netherlands. Five minutes of data were recorded. This data spawned over 50 publications, including two dissertations.

Federal Highway Administration – Next Generation SIMulation

In 2002, the FHWA initiated the Next Generation SIMumlation (NGSIM) project to collect trajectory-level data and enable the theoretical development of new traffic microsimulation behavioral models (Federal Highway Administration Office of Operations, 2017). This was accomplished by synchronizing video cameras, mounted on high buildings near the roadway, that recorded vehicles as they passed through a geofenced area. Post-processing produced position, velocity, and acceleration values at a frequency of 10Hz. There were four roadways that were studied: two freeways and two arterials. The information for the two freeway sites is summarized below in Table 2.1. These datasets are available freely through the Research Data Exchange (Halkias & Colyar, 2006).

Table 2.1 NGSIM Summary Details

Dataset name	I–80	US–101
Site Length (ft)	1650	2100
Number of cameras	7	8
Number vehicles detected	5,684	6,101
Length of data collection (min)	45	45

In 2011, Punzo, Borzacchiello, and Ciuffo observed critical errors within the NGSIM dataset (Punzo et al., 2011). They defined "platoon consistency" and "internal consistency" as metrics capable of capturing the error in a trajectory-level dataset. Platoon consistency was derived from the inter-vehicle spacing between a vehicle pair, while internal consistency was derived from the basic kinematic equations of motion for the following vehicle. The percentage of unrealistic jerk values (> 15 m/s^3), indicative of internal consistency, is between 6% and 17.5% of the NGSIM dataset; additionally, every vehicle in the dataset has at least one jerk value over an acceptable threshold. Moreover, at least 7% of the observed inter-vehicle spacing measurements were unrealistic (i.e., negative inter-vehicle spacing), with maximum bias of platoon consistency as large as 59m. Most importantly, the authors recognized that common filtering techniques applied to reduce the noise in a dataset (e.g., averaging, smoothing, and Kalman filters) do not address the systematic component of errors within the traveled space (Punzo et al., 2011). This led to the need to reconstruct the NGSIM trajectories (Montanino & Punzo, 2013).

In 2015, Montanino and Punzo explored the impact of the errors within the NGSIM data by creating a “traffic informed” method for reconstructing the vehicle trajectories such that the physical (e.g., jerk and speed profiles) and platoon (e.g., inter-vehicle spacing)

integrity are maintained (Montanino & Punzo, 2015). The authors calibrated the Intelligent Driver Model (IDM) car-following algorithm and the Minimizing Overall Braking Induced by Lane Changes (MOBIL) lane-changing algorithm via OptQuest Multistart using both the original and the reconstructed I–80 NGSIM dataset. The selected goodness-of-fit function was the root mean squared error (RMSE) and the selected measure of performance was the inter-vehicle spacing. Surprisingly, these bad trajectories ultimately do not have much of an impact on car-following behavior, but the accuracy of trajectories affects distribution and correlation structure of lane-changing parameters (Montanino & Punzo, 2015).

Drone Data Collection Efforts

The FHWA Office of RD&T recently sponsored a project to explore the impact of lane narrowing as a bottleneck mitigation strategy. As part of this effort, microsimulation models were calibrated with data collected on 10-foot, 11-foot, and 12-foot freeway lanes. Unmanned aerial vehicles, or drones, were utilized to collect trajectory-level data for model calibration. Eight hours of drone and helicopter footage was captured for control sites (i.e., basic segments with 12-foot lanes) and narrowed segments (i.e., basic segments with 10-foot and 11-foot lanes). Data were collected in several cities around the United States:

- Dallas, TX;
- San Antonio, TX;
- Seattle, WA;
- Fort Lauderdale, FL; and
- Honolulu, HI.

Figure 2.1 Aerial still image of Quality Counts drone footage in Honolulu, HI

The trajectories of the vehicles were collected at a frequency of 10Hz. Available data include the following:

- car ID;
- car type (car, heavy vehicle);
- X position (m);
- Y position (m);
- speed (m/s);
- tangential acceleration (m/s^2);

- lateral acceleration (m/s^2); and
- temporal headway (s).

Through use of the DataFromSky (DFS) viewer, the leading vehicle ID can be obtained; this allows the collection of (multiple) leading vehicle position, velocity, and acceleration data with some additional manual effort.

2.2. CAR-FOLLOWING MODELS

This section reviews car-following models in the existing literature. For a more detailed summary of car-following models, see Brackstone and McDonald (1999), Toledo (2007), and Saifuzzaman and Zheng (2014).

2.2.1. Stimulus-Response Models

The most simplistic car-following models are stimulus-response; these models hypothesize that driver reaction—that is, their acceleration—is directly attributed to an external stimulus that was observed. Though the independent stimuli vary between model formulations (e.g., leading vehicle velocity, inter-vehicle spacing), these models are deterministic and assume that the attributes of the lead and following vehicle are known with certainty. An example of a stimulus-response model is the Chandler-Herman-Montroll (CHM) model, which calculates a desired acceleration such that the relative velocity between the following vehicle and leading vehicle goes to zero during steady-state following (Chandler, Herman, and Montroll, 1958).

The original CHM model clearly oversimplified the car-following process. Over the last 70 years, numerous extensions have been made to the model. As an example, the Gazis-Herman-Rothery (GHR) model extends the original CHM model by accounting for

relative distance to the leading vehicle as an explanatory factor in car-following (Gazis, Herman, & Rothery, 1961). Bexelius (1968) expanded the CHM model to account for more than one downstream leader impacting the following vehicle, or what is now known as multi-attention in driving (Bexelius, 1968).

2.2.2. Safety Distance Models

Safety distance models calculate a following vehicle's desired behavior such that the following vehicle can safety react to the leading vehicle should that vehicle decide to come to an abrupt stop. These models are based on Newtonian equations of motion. An example of a safety-distance model is the Gipps model, which calculates the minimum velocity, to be applied after an appropriate reaction time, between two velocity constraints (the first constraint is defined by the driver's desired velocity, while the second is defined by the relationship with the leading vehicle) (Gipps, 1981; Kometani & Sasaki, 1959).

2.2.3. Social Force Models

The IDM was developed as an analytical car-following model that is more robust and able to capture naturalistic driving data. Additionally, it was designed to be able to incorporate traffic flow phenomena, like traffic instability and hysteresis. For IDM, the acceleration behavior is a combination of the force compelling the driver to reach their desired speed and the repellant force encouraging the driver to keep a safe distance from the vehicle it is following; this allows the incorporation of both a free driving and constrained driving condition (Kesting & Treiber, 2008). This model has been expanded to account for multi-attention in driving and is known as the Human Driver Model (HDM) (Treiber, Arne, & Helbing, 2006).

The Modified Field Theory (MFT) model is another type of social force model, where the total driver reaction is the summation of the independent reactions of the driver to external stimuli that have been perceived (Andrew L Berthaume, 2015); using the FHWA Living Laboratory data, this model was extended to many additional forces beyond those accounted for in the IDM, including road signage, pavement markings, and barriers in different types of work zones. The equations developed in this research are designed to be applied to the psychophysical framework developed in Lochrane et al. (2015).

2.2.4. Psychophysical Models

The psychophysical car-following model seeks to explain the natural oscillations in car-following behavior because of the limits of human perception. The cornerstone of psychophysical models is summarized as the following: a car-following phase plane is broken into regimes of reaction (i.e., conscious following) and no reaction (i.e., subconscious following); the division of the plane into regimes is a function of driver "action" points, or points where a driver makes a change in their behavior due to changes in the behavior of the leading vehicle. Drivers react to changes in following distance and relative velocity only when perception thresholds are reached. Each of the regimes defined in a psychophysical framework make different psychological driving assumptions and represent distinctly different driver behavior (e.g., approaching, separating, no reaction) (Olstam & Tapani, 2004).

The roots of psychophysical models can be traced back to work by Michaels (1963), Barbosa (1961), and Todosoiev (1963).

Barbosa (1961) initiated the study of simulator car-following data through the application of phase planes, defined as a Cartesian plane in which different phases (or

regimes) of a physical system can be mapped. In a phase plane, the state variable is placed on the x-axis, while the time derivative is on the y-axis. By proposing following distance as the state variable and relative velocity as the time derivative, Barbosa was the first to visualize the car-following spiral (Barbosa, 1961).

Through his observation that portions of the phase plane trajectory are parabolic—that is, the second derivative of the spacing with respect to the relative speed is piecewise constant—the idea of a "decision point model" was introduced. The decision point model assumes that in close-following conditions, the driver decides to take an action—that is, accelerate—at a constant rate; these actions result in a phase plane trajectory that oscillates around an approximate equilibrium point, where the spacing is optimal and the relative speed is null.

Todosoiev (1963) studied the car-following process with respect to the second order phase plane—that is in the relative acceleration vs. relative speed plane. The parabolic trajectories observed by Barbosa were transformed to rectangular trajectories. Todosoiev coined the term "action points" and defined these points as locations where both the following are true: the acceleration changes magnitude and the algebraic sign of the acceleration changes. The action points and thresholds identified by this definition are points where the effect of the driver's perception become visible through changes in acceleration (Todosoiev, 1963).

Michaels (1963) first showed that drivers respond to changes in the size of the leading vehicle in three separate conditions: simple overtaking, steady-state following, and response to acceleration of a leading vehicle. He is attributed with discovering the importance of perceptual factors and their influence on drivers (Michaels, 1963).

More modern representations of psychophysical car-following models are the ones developed by Wiedemann and Reiter (1992) and Fritzsche (1994). Fritzsche has five self-

define regimes: danger, following 1, following 2, closing in, and free driving (Fritzsche, 1994). The Wiedemann model has four defined regimes: following, closing-in, free driving, and emergency regime (Wiedemann & Reiter, 1992).

2.2.5. Utility Theory-Based Car-Following Models

Li, Dixit, and Hamdar (2016) created a car-following model using expected utility theory from behavioral economics incorporating risk aversion and perception. Constant relative risk aversion was used for risk attitude estimation. Risk perception was successfully calibrated using NGSIM data using nonlinear regression in Strata. The model predicts the driver speed that maximizes utility. The model accurately predicts speed, with $R^2 = 0.90$. They conclude that risk perception is a function of driving experience, risk attitude is not a function of driving experience, and that both risk perception and attitude are function of the leader-follower pair type.

2.2.6. Data-Driven Car-Following Models

Data-driven car-following models are one of the newer categories of car-following models; they apply artificial intelligence and fuzzy logic concepts to imitate the uncertainty and ambiguity associated with human perception and recognition. Neural networks were frequently found in the literature as structures for learning car-following behavior. A neural network is a system of "neurons" that are placed sequentially such that the output of the upstream neuron is the input of the next neuron. The artificial neural network is a machine learning technique inspired by biology and intended to mimic the human brain. Artificial neural networks can model complex systems with many interrelated parameters. Thus, it is a highly logical extension that the car-following decision-making process, which has

traditionally been poorly captured through analytical models, may be well explained using neural networks.

Using IRV data from the VTTI NCDS, a fuzzy rule based neural network was developed to model car-following events; it was calibrated using reinforcement learning (Chong, Abbas, Flintsch, & Higgs, 2013). This effort derived car-following driving rules from NDS data, such that a trained neural network could produce similar behavior. Driver behavior is a function of traffic states; traffic states are continuous, with fuzzy logic applied to partition the states into discrete fuzzy sets. State variables included were spacing, relative velocity, and both the leading and following vehicle's velocity. Reinforcement learning works well for this type of application because it is an agent-based learning that rewards correct decisions and penalizes incorrect decisions, effectively training the agent to respond similarly to the target. First, the authors calibrated two individual drivers using NDS trajectories; later, a basic mega-agent was derived from two drivers' trajectory data. The process is summarized as follows: (i) fuzzy rules scan their associated weights; (ii) optimal actions are selected; and (iii) weights are updated according to a reinforcement learning algorithm. These steps were completed for each timestamp of the event (i.e., at a 10Hz frequency). To avoid reaching local optima, significant training is required; this paper used 400,000 time steps (Chong et al., 2013).

Using two drivers from the NGSIM dataset Hao, Ma, and Xu (2016) created and calibrated an extensive fuzzy-rule based car-following model. The authors used a genetic algorithm to calibrate the IDM as a base case; the goodness-of-fit function was mean squared error (MSE) between simulated and actual velocity. The authors calibrated the model using three drivers from the NGSIM dataset; model performance was validated using two different drivers from the same dataset. Their model replaces the stimulus-response framework with a five-layer structure: perception, anticipation, inference,

strategy, and action. This included eight fuzzy sets, 152 fuzzy rules, and 80 attributes. In the perception stage, the neural network collects information by imitating observations of variables; this observation may be deterministic (e.g., follower speed, number of lanes) or fuzzified (e.g., lead vehicle speed, space headway). In the anticipation stage, the neural network predicts the acceleration and velocity of the leading vehicle in the short-term. In the inference stage, the neural network imitates the decision-making process of the human; it generates multiple "reasonable control instructions" based on perception and anticipation. In the strategy state, the algorithm selects the most optimal control instruction as a function of safety, degree of comfort, and degree of satisfaction. In the action stage, the driver reacts based on what was learned in the perception, anticipation, and strategy states. The authors found that their fuzzy set car-following model performs significantly better than the IDM because it analyzes information acquisition and the decision-making process (Hao et al., 2016).

Panwai and Dia (2005) used IRV data collected by the Robert Bosch GmbH Research Group. Relative speed and relative space headway were used to predict the following vehicle's speed at each time step. Several artificial neural network architectures, learning rules, and transfer functions were studied during this effort. The architectures explored included back-propagation, fuzzy predictive adaptive resonance theory (ARTMAP), and radial basis function network. All models resulted in classification rates exceeding 96% accuracy. Additionally, both microscopic and macroscopic data was used to verify the efficacy of using artificial neural networks to predict driver behavior (Panwai & Dia, 2005a).

Khodayari, Ghaffari, Kazemi, and Braunstingl (2012) applied a neural network to NGSIM data; they used 6101 vehicle trajectories recorded over the course of 45 minutes on eastbound US–101. They did not use the reconstructed NGSIM data, electing to filter

the data with a moving average filter with a window of 1s. The estimated instantaneous reaction delay, relative speed, relative distance, and following vehicle velocity were used as inputs to the artificial neural network; the following vehicle acceleration was the output. A back-propagation algorithm was used to train the model (Khodayari et al., 2012).

2.3. DRIVER HETEROGENEITY

There are two types of driver behavior heterogeneity: intra-driver heterogeneity and inter-driver heterogeneity. Intra-driver heterogeneity describes the phenomenon where a single driver behaves inconsistently within themselves while driving. Instantaneous level of congestion (e.g., congested vs. uncongested condition), operational conditions (e.g., presence of a work zone), and asymmetries between a driver's behavior while accelerating vs. decelerating are all examples of intra-driver heterogeneity. Conversely, inter-driver heterogeneity describes the phenomena where different drivers drive differently under the same conditions. Differences in desired speed, following distance, reaction time, and acceleration/deceleration rates are empirical examples of inter-driver heterogeneity. An extensive literature review has shown that though the existence of both types of driver heterogeneity is quite strongly supported, little work has been conducted to understand the root causes of driver behavior heterogeneity or how to accurately capture it in car-following models.

2.3.1. Heterogeneity in Empirically Collected Data

In response to Brackstone and McDonald's car-following model review from an engineering perspective, Ranney (1999) wrote a summary of the psychological factors that influence car-following models. Ranney hypothesizes that the individual factors that

influence car-following behavior can be broken into two categories: individual differences and situational factors. The relative influence of these two factors is hypothesized to be a function of the level-of-service of traffic flow near the subject vehicle. He hypothesized that at lower densities, individual differences are more influential; at higher densities, situational factors have more influence. Individual differences suggested by the author include age, gender, marital status, vehicle type, and risk-taking propensity; suggested situational factors include time of day, weather, and road type (Ranney, 1999). Section 2.3.1 discusses heterogeneity observed in empirical data. This chapter is categorized into subsections by the surrogate for driver behavior (e.g., gap, reaction time)

2.3.1.1. Heterogeneity in Desired Gap and Headway Selection

Of all the surrogates for driving behavior encountered in articles reviewed for Section 2.3.1, the most commonly observed surrogate was observed gap and headway selection. Thus, this chapter is further segmented into heterogeneity in gap/headway as a function of driver specific attributes and as a function of the situational environmental factors.

Heterogeneity attributed to driver attributes

Evans and Wasielewski (1983) used a photographic technique to record observations of people's naturalistic headway selection. Over 12,000 observations of headway were recorded. Evans and Wasielewski found that both gender and age have an impact on headway choice. They observed that females chose longer headways than males. Additionally, desired headway was positively correlated with driver age. Lastly, it was observed that the age of the vehicle, which is possibly an indirect indicator of driver

income, was correlated with decreased headways. These relationships were all statistically significant at the 5% level (Evans & Wasielewski, 1983).

Elander, West, and French (1993) observed that younger drivers tend to maintain shorter headways. This research studied empirical data and found that both driving skill and driving style contribute to crash risk. This research clearly indicates that young drivers generally adopt more aggressive/risky driving styles and older drivers tend to be more cautious than middle age drivers (Elander et al., 1993).

A study by Boyce and Geller (2002) observed that young age, defined as between 18 and 25, is a strong predictor of following distance. An IRV was used to collect data from 61 drivers between the ages of 18 and 82. Older drivers maintained a safe vehicle speed more than both younger and middle-aged drivers. Men were observed to follow at an unsafe time gap more frequently than women (Boyce & Geller, 2002).

Wu, Brackstone, and McDonald (2003) further corroborated the observation that driver age impacts headway selection. They observed that drivers over the age of 59 selected an average headway of 1.83s; this is 23% higher than a driver between the ages of 23 and 37 (Wu et al., 2003).

Underwood (2013) explored how driving behavior changes with experience as a function of age. The drivers were between 17–19 and 22–44 years old. Drivers were tested in an IRV zero, three, and six months after acquiring their license. It was observed that all drivers tended to increase their average speed as they gained experience driving. The older group showed stronger indications of more conservative driving with experience, as their headways increased over time. This study concluded that there is a correlation between age and driving experience with regard to driving style (Underwood, 2013).

Lochrane, Al-Deek, Jiang, Dailey, and Shurbutt (2017) showed that when driving through work zones driver experience is the most critical factor in identifying a driver's

average time gap; drivers with less than 10 years of experience followed at significantly shorter gaps than the rest of the population (2.65s vs. 3.28s). In non-work zones, driver experience was once again the most significant predictor of mean time gap; drivers with less than 10 years driving experience selected a shorter time gap. For drivers with less driving experience, the next most significant indicator was educational attainment; surprisingly, those with an advanced degree selected an average mean time gap of 1.52s, compared to 2.12s observed in the subpopulation with a college degree or less. For more experienced drivers, gender was the next most significant indicator; males selected a shorter average time gap, almost a full second shorter than the average female time gap. The data for this study was collected via IRV in FHWA's Living Laboratory. The data from 12 drivers was reduced to over 5,000 logged time gap points for the decision tree analysis (Lochrane et al., 2017).

Heterogeneity attributed to traffic conditions and leading vehicle type

Dijker, Bovy, and Vermijs (1998) examined the differences in headways maintained by different vehicle type categories on two Dutch freeways. Holding speed fixed, they found significant differences in car-following headways between congested and uncongested conditions. They observed that for passenger cars, the distance gap in congested conditions is smaller than the distance gap observed in uncongested conditions. However, for heavy vehicles the distance gap was approximately the same regardless of the traffic flow (Dijker et al., 1998). Brackstone, Waterson, and McDonald (2009) found conflicting results: they concluded that the prevailing flow of traffic has negligible impact on headways. For this analysis, data was collected via IRV equipped with sensors capable of collecting following vehicle speed and relative distance to the leading vehicle. Other conclusions drawn from this study include that driver behavior is strongly impacted by lead

vehicle type, driver behavior is not influenced by roadway type, and that intra-driver variation in following driver behavior—including reaction time, variation in lateral position, and variation in speed—exist in the data (Brackstone et al., 2009).

Yin et al. (2009) found that in uncongested conditions, vehicle headway distributions are best described by a lognormal distribution; in congested conditions, they are best fitted by a log-logistic distribution. The data for this analysis was obtained from video recordings of two busy expressways in Beijing, China. Kolmogorov–Smirnov (KS) tests were performed to confirm the model fits were statistically significant (Yin et al., 2009).

Zhu, Wang, and Wang (2016) explored the inter- and intra-driver heterogeneity observed in naturalistic headway selection as a function of operating speed, weather, roadway type, and traffic density using naturalistic data collected in China; a single IRV was driven by 55 drivers for two months, resulting in 1489 car-following events. The primary conclusions were that the distribution of car-following headways is most closely represented by the lognormal distribution. Additionally, drivers tended to select longer headways when speeds are reduced, during low light conditions, and when traffic is congested. Three levels of congestion were examined: sparse, moderate, and high; these were classified as a function of traffic density. The alteration of driver behavior was statistically significant, with p-values < 0.01, for the aforementioned independent variables (Zhu et al., 2016).

Geng, Liang, Xu, and Yu (2016) applied the SHRP2 NDS data to determine general differences in car-following behaviors under different driving conditions (e.g., traffic, vehicle types, weather, etc.). They determined that the average headway for each event is a function of traffic condition, lead vehicle type, time of day, and weather at a statistically significant level (Geng et al., 2016).

Ye and Zhang (2009) explored the differences in headway between different leader-follower vehicle pairs. They found that the passenger car / passenger car (PC-PC) pair had the shortest headway, while the heavy vehicle / heavy vehicle (HV-HV) pair had the longest headway. This, however, varied with level of congestion. In uncongested traffic, the headway distribution between all four combinations of leader-follower pairs were fairly consistent. However, in congested traffic, the headways tend to be smaller for all leader-follower vehicle pairs; additionally, the dispersion of headway within a leader-follower pair category is significantly higher under congested conditions (Ye & Zhang, 2009).

Weng, Meng, and Fwa (2014) segmented vehicle headway data collected in work zones on expressways in Singapore into four categories by leader-follower vehicle pairs: HV-HV, PC-PC, HV-PC and PC-HV. Using a non-parametric analysis of variance (ANOVA) and a Mann-Whitney test, they confirmed that the median values of the four categories of vehicle-pair headways varied significantly as a function of level of congestion, percentage of heavy vehicles, work intensity, and lane position. Additionally, using maximum likelihood estimation and KS test techniques, they concluded that the headway distribution when a PC was the leading vehicle was best described by a log-normal distribution, while the headway distribution when a HV was the leading vehicle was best described by an inverse Gaussian distribution (Weng et al., 2014).

Aghabayk, Sarvi, Forouzideh, and Young (2013) used the NGSIM data to explore the inter-driver heterogeneity in spacing and acceleration distributions. This research looked at variability in driver behavior as a function of leading vehicle type. Four types of pairs were considered: PC-PC, PC-HV, HV-PC, and HV-HV. Using space-time diagrams, the authors first observed different headways as a function of leader-follower pair. Next, the authors categorized the acceleration data of the following vehicle in $0.1 m/s^2$ bins and created a distribution. Kolmogorov-Smirnov tests were conducted to compare the

discretized acceleration behavior as a function of leader-follower pair. This analysis confirmed that acceleration distributions are different at a statistically significant level between the leader-follower pair categories. This effort also applied a Cochrane-Orcutt procedure to determine the impact of independent variables on acceleration (in absence of the issue of autocorrelation). This analysis indicates that relative velocity is the most important variable for all four vehicle pairs. The PC-PC pair had the most variables identified as important (four), while the HV-HV pair had the least (two); this indicates that the PC-PC leader-follower pair is the most sensitive to external stimuli while HV-HV leader-follower pair is the least. Desired spacing is only identified as an important stimulus in the PC-PC pairing (Aghabayk, Sarvi, & Young, 2014).

2.3.1.2. Heterogeneity in Desired Speed

Elander, West, and French (1993) observed that younger drivers travel at faster speeds than older drivers. This research studied empirical data and found that both driving skill and driving style contribute to crash risk. This research clearly indicates that young drivers generally adopt more aggressive/risky driving styles and older drivers tend to be more cautious than average (Elander et al., 1993). De Waard, Dijksterhuis, & Brookhuis (2009) observed similar results via a human driving simulator; they observed that older drivers (65+) kept a lower speed than younger drivers during a merging scenario.

A study by Boyce and Geller (2002) observed that young age, defined as between 18 and 25, is a strong predictor of risky behavior, such as speeding. An IRV was used to collect data from 61 drivers between the ages of 18 and 82. Driver age and type A personality characteristics were significant predictors of vehicle speed and following distance. Men and women were observed to have an equivalent propensity towards

speeding (Boyce & Geller, 2002). The absence of gender as a predictor of risky speed selection was also observed by Ericsson (2000).

2.3.1.3. Heterogeneity in Acceleration Behavior

Ericsson (2000) found that men tended to accelerate harder than females, especially on local feeder roads in a residential area. Additionally, she found that deceleration rates were lower during peak hour compared to off-peak periods, possibly indicative of intra-driver heterogeneity as a function of prevailing traffic conditions. This study was intended to improve emission models of urban traffic (Ericsson, 2000).

Wang, Wang, Chen, and Jing (2011) observed intra-driver heterogeneity between the acceleration and deceleration phases of a single driver. Using data from Dutch motorways, they observed that 65% of drivers demonstrate different driving styles between the two states. The heterogeneity is so abundant that a car-following model calibrated for the acceleration behavior of a driver cannot sufficiently capture the deceleration behavior of that same driver. They recommend a multi-phase car-following model to better capture this heterogeneity (Wang et al., 2011). Furthermore, Li and Chen (2017) concluded that the strong right-skew of empirical headway distributions are a byproduct of intra-driver heterogeneity between the acceleration and deceleration process; this has far reaching inferences on macroscopic traffic flow phenomenon (Li & Chen, 2017).

Itkonen et al. (2017) explored the relationships between time headway, acceleration, and jerk in a simulated environment where vehicle and environment were controlled variables. They observed trade-offs made between “close but jerky” vs. “far but smooth” following behavior; they hypothesize that this tradeoff serves as an indicator of a possible latent factor underlying the driving styles hypothesized in traffic psychology

literature. Their highway driving simulator (HDS) experiment consisted of 15 participants (five males, ten females) with a mean age of 31. The authors analyzed the geometric mean of time headway and the mean of the absolute values of acceleration and jerk. They found that, on average, drivers fall into one of two categories: they decide to match the leader's speed at close distances (with stronger reactions when the lead vehicle accelerates/decelerates) or they maintain a longer gap and regulate their speed in a more "calm" manner. They proposed that the relationship between short time gaps and the "jerky" driving tradeoff can be interpreted as an "intensity-calmness" parameter of driving style (Itkonen et al., 2017).

Berthaume, James, Hammit, Foreman, and Melson (2018) observed heterogeneity in acceleration behavior between different road types (i.e., highway, freeway), operational conditions (e.g., no work zone, advanced warning zone, work zone with shoulder closure, and work zone with lane closure), and levels of congestion (i.e., congested, uncongested). They plotted trajectory-level data on a psychophysical plane (i.e., relative speed vs. relative spacing) and used statistical tests to show the statistically significant differences in driving behavior between the different conditions.

2.3.1.4. Heterogeneity in Reaction/Braking Time

Warshawsky-Livne and Shinar (2002) analyzed the impact of age, gender, vehicle transmission type, and event uncertainty on the observed braking time. Braking time was split into components of perception-reaction time and brake movement time. Perception-reaction time was positively correlated with age at a statistically significant level; age was not found to be an explanatory factor of brake movement time. The reaction time of the older group of drivers (mean age 62) was 20% higher than that of the youngest group of

drivers (mean age 23). Conversely, gender was found not to be an explanatory factor of perception-reaction time; however, the average brake-movement time for male drivers was slightly longer than that for female drivers. This study was completed in a driving simulator, with data collected for 72 subjects (Warshawsky-Livne & Shinar, 2002).

Mehmood and Easa (2009) developed a driver reaction time model in a car-following context accounting for various statistically significant human factors. Driver's age, gender, speed, and spacing were all found to be statistically significant in predicting the acceleration/deceleration reaction time. The data for this study was collected via driving simulator for sixty participants between the ages of 18 and 70. This effort observed that reaction time for both acceleration and deceleration decrease with age; they also observed that females react more slowly than males (Mehmood & Easa, 2009).

2.3.1.5. Heterogeneity in Traffic Violations and Crashes

Laapotti, Keskinen, and Rajalin (2003) analyzed younger drivers' attitudes and self-reported driving behavior. They showed that between the years of 1978 and 2001 the gender difference in traffic offenses has remained steady, with females committing fewer traffic infractions and having a lower crash rate (Laapotti et al., 2003).

Corbett (2007) observed that there is a gender gap between males and females in terms of car-related crimes and convictions (e.g., speeding); she also observed that females are more heterogeneous with their driving styles, as there is a "ladette" group of younger females whose driving styles are more reflective of young males (Corbett, 2007).

2.3.1.6. Personality Questionnaires and Driving Behavior

In Brackstone (2003), data was collected via IRV for 11 participants (ten males, one female). The possible correlation between the Sensation Seeking and Internality-Externality Scales was explored with naturalistically collected following distance. A positive correlation was observed between driver externality rating and low speed behavior (30mph or less), whilst a negative correlation was observed between sensation seeking scale V and low speed behavior. For higher speeds, these scales are not valid predictors; instead, the participant's rating of aggressiveness is most strongly correlated with driver behavior (Brackstone, 2003).

Ulleberg and Rundmo (2003) distributed a questionnaire designed to collected data to study the importance of personality traits and their relationships with risk behavior in traffic. High scores on sensation seeking, normlessness, and aggression were associated with risk-taking attitudes and risky driving behavior. Individuals scoring high on altruism and anxiety traits were less likely to report risky behavior in traffic. The personality variables accounted for 47% of the total observed attitude variance; this supports the hypothesis that risk-taking attitudes could be explanatory variables for the variance in driving behavior. The results of this study strongly support the importance of exploring the indirect effects of personality relative to risk-taking attitudes and driver behavior. The results for this study were derived from the responses from 1932 adolescents in Norway (Ulleberg & Rundmo, 2003).

Ishibashi, Okuwa, Doi, and Akamatsu (2007) collected data from sixteen drivers between the ages of 22 and 52, equally split between male and female, via IRV; the participant's driving style was collected through a driving style questionnaire (DSQ). The researchers were interested in low-speed car-following events, defined as 4–40 kph follower speed and less than 50m following distance. They collected data on drivers'

acceleration and deceleration behavior. It was observed through multiple regression analyses that the DSQ indices are statistically significant explanatory variables for following distance (Ishibashi et al., 2007).

Kleisen (2011) used the multi-dimensional driving style inventory (MDSI) to compare driving styles between younger drivers of both genders; it was observed that females scored higher with respect to positive driving styles–identifying as "patient" and "careful"—whereas males were identified as "risky", "angry", and "high-velocity" (Kleisen, 2011).

2.3.1.7. Heterogeneity Observed Through Data-Driven Clustering Techniques

Constantinescu, Marinoiu, and Vladoiu (2010) applied (i) hierarchical clustering techniques based on Euclidean distance and Ward's method and (ii) principal component analysis from exploratory statistics to establish relationships with driving behavior. The data from 23 different drivers in a single city, Bucharest, was used. This research used driving parameters, not driver characteristics, to cluster the data by driving style. Percent of time with speed in excess of 60 kph, average speed, standard deviation of speed, standard deviation of acceleration, mean and standard deviation of positive acceleration values, and mean and standard deviation of negative acceleration (braking) values were mined from the data as explanatory variables. Data was sampled at a frequency of 1Hz. Aggressiveness, ranging from moderately low to high, and speed, ranging from low-moderate to high, were used as principal components. Positive acceleration, ranging from moderate to high, and braking, ranging from smooth-moderate to sudden, were analyzed as rotated components. This resulted in six observed clusters of behavior for this dataset. As future research, they recommend accounting for individual characteristics when developing the clusters of driving styles (Constantinescu et al., 2010).

Wu, Du, Qi, and Xu (2015) applied clustering techniques to extract latent driving states from longitudinal driving behavior for ten drivers; the data was collected via an IRV. They applied an ensemble clustering method combining the kernel fuzzy C-means algorithm and the modified latent Dirichlet allocation model to determine underlying structures for 'aggressive', 'cautious', and 'moderate' driving styles. The clustering method was based on longitudinal driving data (e.g., acceleration, headway) and not characteristics of the drivers (Wu et al., 2015).

Fernandez and Ito (2016) developed a fuzzy rule-based system to classify drivers into different driver profiles based on their behavior. Inputs to the fuzzy model included the percentage of time the driver used the gas/brake pedals and the speed of the vehicle. This work also accounted for driver age using trapezoidal membership functions. The rule-based system either classifies the driver as "aggressive" or "passive". Ultimately, the system is used to predict route choice and time to destination in a small example with five roads and six intersections (Fernandez & Ito, 2016).

Li, Li, Cheng, and Green (2017) estimated the driving style of 28 drivers as a function of maneuver transitions. A conditional likelihood maximization method was applied to extract maneuver transition patterns from naturalistic driving data; a random forest algorithm was then applied to classify driving styles. Transitions between five maneuver states—free driving, approaching, near following, constrained left lane changes, and constrained right lane changes—were found to reliably classify drivers into different driving styles—low-,moderate-, and high-risk. This effort did not consider any driver characteristics or the prevailing congestion level as explanatory variables for driving style. This data was only collected during light traffic conditions (LOS A) (G. Li et al., 2017).

2.3.2. Capturing Heterogeneity in Car-Following Models

This chapter discusses the limited research in establishing and accounting for heterogeneity in car-following models. The existence of heterogeneity is strongly supported in the literature. However, only level of congestion and leading vehicle type have been used as explanatory factors for driver heterogeneity in car-following models.

2.3.2.1. Intra-Driver Heterogeneity

Wang, Wang, Chen, and Jing (2010) leveraged the NGSIM data to study differences in intra-driver behavior during the acceleration and deceleration process using a Helly, IDM, and Gipps model. The authors found that Helly and IDM predict driving behavior better than Gipps. Additionally, drivers respond to changes in traffic more quickly and intensely in the deceleration process than in acceleration. The response to speed difference stimuli is very different between the acceleration and deceleration phases; conversely, the response to distance gap stimuli is more equivalent between the acceleration and deceleration phases. More than 65% of drivers were observed to drive according to different optimal car-following models in the acceleration and deceleration process. The optimal model parameters of the acceleration process were observed to be very different from that of the deceleration process. They concluded that the IDM was the most robust model. Helly and Gipps were both observed to be highly sensitive to time delay. Lastly, Gipps was observed to have better consistency between the acceleration and deceleration phases, but performed the worst overall (Wang et al., 2011).

Zheng, Suzuki, and Fujita (2012) used three pairs of vehicle trajectories from NGSIM dataset to calibrate several car-following models. One trajectory collected during congested stop-and-go traffic was used for calibration, while two separate trajectories, one from more severe congestion and one collected in uncongested conditions, were used for

validation; the trajectories were collected from different drivers. The authors calibrated five different car-following models: GHR stimulus-response model, Gipps safety distance model, Newell's trajectory translation model, Cellular automa, and the optimal velocity model (OVM). The authors calibrated the model using a genetic algorithm seeking to minimize the relative error between predicted and observed relative spacing. They observed high error rates (i.e., 17–34%) even though the parameters were all well within the range of suggested values. They concluded that these models suffered from overfitting in the validation process; however, it is possible they failed to recognize intra-driver heterogeneity induced by the vastly different conditions under which their calibration and validation datasets were collected (Zheng et al., 2012).

Abbas, Higgs, and Medina (2011) used the VTTI 100-Car NDS data. Ten car-following periods from five different drivers were used in the calibration process. A car-following framework was created such that the car-following regimes were defined by the traditional Wiedemann model, but the equations for acceleration in each regime were replaced by the stimulus-response GHR model, with different parameters selected for each of the approaching, closely approaching, acceleration following, and deceleration following regimes. The calibration process was completed with a genetic algorithm. The primary result is that this hybrid model is more independent of the behavior of the lead vehicle; this is in direct contrast with most other car-following models, where the lead vehicle behavior is the most influential attribute (Abbas, Higgs, & Medina, 2011; Abbas, Higgs, Medina, & Yang, 2010).

Papathansopoulou and Antonio (2016) created a flexible car-following model using ten vehicle trajectories from the reconstructed NGSIM dataset by accounting for the density of the adjacent lanes. The objective of the calibration problem was to minimize the normalized RMSE between modeled and observed speed. This model was calibrated using

a pair of vehicle trajectories and validated using nine pairs of vehicle trajectories. A "metamodel" was developed to evaluate the magnitude of the effect of the independent variables. The authors found that in lower speed, higher density conditions, car-following behavior is conservative and harder to model accurately. Accounting for density helps this new data-driven model perform better than a calibrated Gipps model (Papathanasopoulou & Antoniou, 2016).

Hammit, Ghasemzadeh, James, Ahmed, and Young (2018) and Hammit, James, Ahmed, and Young (2019) observed intra-driver heterogeneity between driving behavior in clear and adverse weather conditions using a sample of the SHRP2 NDS dataset. They used matching trips (i.e., same driver, same road, nearly same time of day) during clear and adverse weather conditions (i.e., light rain, moderate rain, heavy rain, fog, and snow) for their analysis. They showed that these differences in driving behavior can be captured and modeled in microsimulation using the Gipps and Wiedemann 99 models, respectively.

2.3.2.2. Inter-Driver Heterogeneity

Brockfeld, Kuhne, and Wagner (2004) calibrated ten different car-following models using IRV data from an eight-car platoon on a test track collected in Japan in 2001. The downhill simplex method was used to minimize the percent error and absolute error between simulated and empirical time headways. The authors found errors of 9–24% for each model. No model significantly outperformed the rest for all drivers, with average errors between 15.1% and 16.2%. Using holdback data, the validation error was between 17 and 22%, which they attributed to overfitting the data. The differences between individual drivers were found to be more significantly different than the differences between models (Brockfeld et al., 2004).

Ranjitkar, Nakatsuji, and Asano (2004) calibrated the Gipps, Krauss, Newell, Castello, Brando, and Excess Critical Speed stimulus-response (ECS) model using 2001 IRV data from an eight-car platoon on a test track in Japan. The researchers applied a genetic algorithm to minimize the percent error between simulated and observed velocity and headway. Velocity was found to be the better measure of performance, with errors between 3.87% and 4.71%, compared to 12.04–12.91% for headway. They concluded the optimal parameters inter-driver variation is much more significant than the differences between model variables (Ranjitkar et al., 2004).

Punzo and Simonelli (2005) used the smoothed Naples IRV data to calibrate four models: MITSIM, Gipps, Newell, and the IDM. The authors calibrated the models by minimizing the root mean square percentage error (RMSPE) between predicted and observed headway, speed, and spacing. They found that although it is easier to calibrate to speed (i.e., a lower RMSPE is achievable), it is much more robust to calibrate based on spacing because the error accurately accumulates through the trajectory. However, they observed that though the calibration errors (mean error of 15.5%) were on par with those noted in the literature, cross-validation completed on withheld data resulted in much higher errors (mean error of 22.31%), which they attributed to overfitting (Punzo & Simonelli, 2005).

Ossen and Hoogendoorn (2005) used the helicopter trajectory data, whose collection was sponsored by TNO, to calibrate three different functional formats of the GHR stimulus-response model. They estimated the parameter ‘C’ via least squares fit for all feasible reaction times; if the relationship between the stimuli and response was statistically significant at the 5% alpha level, the estimation of C as a function of reaction time was kept for further analysis. The optimal reaction time was found using a Bayesian regulation objective function to reduce the likelihood of overestimating extreme values of

reaction time. They discovered that inter-driver heterogeneity is existent in the trajectory data, with variation observed with both optimal parameters and the car-following functional forms between the different leader-follower pairs. They were able to establish a relationship between the stimuli (i.e., relative speed, relative distance, and speed of following car) and the response (acceleration of the following car) for 80% of the leader-follower pairs. They determined that different drivers have varied reactions to different stimuli and that drivers that react to the same stimuli do so with different sensitivities (Ossen & Hoogendoorn, 2005).

Ossen, Hoogendoorn, and Gorte (2006) explored the heterogeneity in the aerially collected data funded by the TNO. The authors calibrated stimulus-response (i.e., GHM, Bexelius, Helly, Addison and Low, and Tampère), safety distance (i.e., Gipps), and trajectory translation (i.e., Newell) models. The authors calibrated the parameters of each model assuming the reaction time is known or given. They applied simplex search with multistart to solve an optimization problem seeking to minimize Theil's U with respect to relative speed and relative distance. The calibration process was completed for a range of feasible reaction times (i.e., 0.5–3.5s). The reaction time yielding the optimal objective value was selected. The models were evaluated on two levels: average performance of each model and on the level of the individual driver. They analyzed the empirical cumulative density function (CDF) of the minimized objective values, the change in optimal value of objective function when model B is used instead of model A, and the parameter empirical CDFs. The simplest models were not able to capture the dynamics of car-following behavior correctly. For more complex models, optimal parameter settings differ among drivers. They determined inter-driver differences cannot be captured via different parameter settings alone (i.e., analysis indicated that the driving styles of different drivers

appear inherently different; in other words, different car-following models are needed to model a population) (Ossen et al., 2006).

Brockfeld and Wagner (2006) used four different sets of data to explore driver heterogeneity: (i) data collected by Daganzo in 2000 consisting of travel times between points on a single lane road; (ii) an eleven car platoon on a test track collected in 2013; (iii) data collected on I–80 using loop detectors with flow, occupancy, and speed aggregated to 15s averages; and (iv) NGSIM data, 1 km in length for 45 minutes collected at a frequency of 10Hz. They calibrated a total of thirteen models. They found that the average calibration error was about 15% (range: 12–20%) and the validation error is on average 20% (range: 15–25%). They concluded that almost all models perform similarly and that simple models perform no more poorly than the most complicated ones. Ultimately, they found the difference between models is smaller than the difference between different drivers, again underscoring the importance of understanding and accounting for driver heterogeneity (Brockfeld & Wagner, 2006).

Chen, Li, Hu, and Geng (2010) applied the I–80 NGSIM data to calibrate the MITSIM and IDM car-following models. The authors concluded that neither MITSIM nor IDM can accurately simulate all driving behaviors. They observed that the calibrated parameters approximately form a hyperplane. Within the hyperplane, there is little clustering; instead, points are dispersed, which can be interpreted as inter-driver heterogeneity (Chen et al., 2010).

Sangster, Rakha, and Du (2013) applied the VTTI 100-Car NDS data to study inter-driver heterogeneity. Two thousand car-following events, recorded by eight drivers traversing a multilane highway, were extracted for approximately 1,000 hours of car-following. Four different car-following models were calibrated: Gipps, IDM, GHR, and Rakha-Pasumarthy-Adjerid (RPA) models. The models were calibrated using a variant of

the RMSPE metric with speed and space headway as the measures of effectiveness; the optimization problem was solved with the evolutionary non-linear approach in Excel. The authors found that the RPA model best simulates the trajectory-level data, followed by the Gipps model. However, the RPA and Gipps model both produce less variability in behavior when plotted as a fundamental diagram compared to the observed data (Sangster et al., 2013).

2.3.2.3. Combined Inter- and Intra-Driver Heterogeneity

Soria and Elefteriadou (2011) applied IRV data collected by the Florida Research Center to calibrate the Gipps, Pitt, MITSIM, and modified Pitt model using Solver in Microsoft Excel. This paper assessed car-following models and their performance under different conditions and for different driver types; the goal of this paper was to improve the application and understanding of existing car-following models. Driver types were categorized into different aggressiveness categories as a function of the number of discretionary lane changes and the observed speed when driving under free-flow conditions. The models were calibrated to minimize RMSE between observed and simulated speed and spacing. The author concluded that it is more robust to calibrate based on spacing. They also found that congested conditions are easier to calibrate than uncongested conditions; additionally, the best performing model varied between congested (Gipps, MITSIM) and uncongested conditions (MITSIM). Analysis of the calibrated models indicated that under congested conditions driver behavior is better predicted when calibrated by level of congestion, while in uncongested conditions the performance is better when calibrated by driver type. Additionally, it is shown that conservative and aggressive driver behavior is more accurately captured when calibrated by driver type. In other words,

under congested conditions intra-driver heterogeneity as a function of traffic density is the most explanatory factor of driving behavior. Conversely, in uncongested condition, inter-driver heterogeneity as a function of aggressiveness level is most critical; this may be attributable to the fact that driver aggression, as defined by the author, is a function of speed under free flow conditions and discretionary lane changes (neither which can be observed accurately in congested conditions) rather than driver attributes (Soria, 2010; Soria & Elefteriadou, 2011).

Ossen and Hoogendoorn (2011) explored the impact of leading-vehicle type as an explanatory factor for inter-driver heterogeneity. The authors calibrated eight different car-following models: CHM, Bexelius, Tampere, Addison and Low, Gipps, IDM, OVM, and the two leader Lenz models; they used the helicopter data collected by the Technical University of Delft funded by the TNO. A constrained non-linear optimization algorithm based on the simplex method was used to minimize Theil's U considering both follower speed and relative distance. Six hypotheses regarding driver heterogeneity were tested: three hypotheses explaining driving style heterogeneity and three hypotheses regarding heterogeneity within a driving style. For the extraction of a car-following event, the authors required that no lane changes take place for a triple of vehicles whilst the data was collected, there was a speed change of at least 5 meters per second during the data collection, and the cars were adjacent for at least 15s. From this effort, the authors confirmed that different drivers follow different styles. The Tampère model was most optimal for 34% of drivers; the Lenz performed best for 27% of the remaining drivers. Moreover, the standard deviation of error of the models further suggested significant heterogeneity within the data. Additionally, it was shown that within the passenger car subpopulation there was a significant amount of unexplained heterogeneity. Conversely, HVs were much more robust than the passenger cars; they tend to travel with a more

constant speed and are less eager to return to their desired space gap. Lastly, they show that characteristics of a facility (e.g., capacity) are highly dependent on the longitudinal driving characteristics of the vehicles on the facility. This further underscores the importance of accurately characterizing driver heterogeneity in microscopic models (Ossen & Hoogendoorn, 2011).

Higgs, Abbas, and Medina (2011) calibrated the Wiedemann model for different drivers segmented into different speed ranges using the VTTI 100-Car dataset. A genetic algorithm was used to solve the calibration problem, which sought to minimize the difference between model and observed following vehicle speed. They found that the Wiedemann coefficients were vastly different for different speed bins and that the variation seemed to be driver dependent (Higgs, Abbas, & Medina, 2011).

Trieber and Kesting (2013) documented one of the most extensive studies regarding the viability and efficacy of using microscopic trajectory data to calibrate car-following models. In this effort, they observed the existence of both inter- and intra-driver heterogeneity. Using aerial and floating car data to calibrate the IDM, OVM, and the full difference velocity (FVD) models, they determined that different models represent different driving styles more appropriately; this supports the existence of inter-driver heterogeneity. Additionally, they studied the importance of different parameters representing different driving situations (e.g., cruising, approaching, following, free-driving, etc.). They concluded that models that use different calibrated parameters to capture different driving situations are the most robust (Treiber & Kesting, 2013b).

Higgs and Abbas (2013) explored differences in driving styles using data of three drivers from the VTTI 100-Car study; these drivers were selected because they represented a high-risk, medium-risk, and low-risk driver. Risk was assessed by the relative number of "conflicts" experienced by each driver while data was collected. Ten car-following periods

were used for each driver. The data was segmented using time-dependent clustering, which identified the optimal segment lengths; each segment has its own centroid. Each centroid has eight dimensions: longitudinal acceleration, lateral acceleration, yaw rate, vehicle speed, lane offset, yaw angle, range, and range rate. Afterwards, the segments were clustered to find similar behaviors in the data via a k-means clustering algorithm. This analysis shows that common car-following model assumptions regarding heterogeneity (i.e., not accounting for inter-driver and intra-driver heterogeneity) are strongly violated in the naturalistic driving study data, as behavior varied between drivers, between car-following periods, and within a single car-following period (Higgs & Abbas, 2013).

Higgs and Abbas (2014) applied the VTTI 100-Car NDS datasets to study driving behavior heterogeneity. They used ten different drivers and over 3000 car-following periods. This research calibrated the GHR, IDM, velocity difference, Wiedemann, hybrid Wiedemann-GHR, and clustered GHR car-following models. This research was conducted at four different resolutions: (i) all drivers (i.e., all 3000+ car-following instances from the ten drivers are calibrated to obtain one set of parameters), (ii) driver specific (i.e., one set of calibrated parameters is obtained for an individual driver using all of the car-following periods for that driver), (iii) car-following period specific (i.e., one set of calibrated parameters is obtained for each car-following period), and (iv) cluster specific (i.e., one set of calibrated parameters is obtained for each cluster of data in a car-following period). Calibration parameters were obtained by minimizing the RMSE between the simulated and observed speed. At the driver specific level, the hybrid-Wiedemann and the velocity difference models reported the lowest error. At the car-following period specific level, the hybrid Wiedemann model performed the best. At the segmented car-following period level, the segmented GHR model outperformed the rest of the models. Additionally, at the car-following period level, the segmented IDM and segmented GHR models performed better

than their unsegmented counterparts; however, the segmented velocity difference model performed significantly worse than the original velocity difference model. This highlights the advantages and disadvantages of different models at different levels of analysis (Higgs & Abbas, 2014).

Higgs and Abbas (2015) applied the VTTI 100-Car and truck NDS datasets to study inter-driver differences between PC and HV drivers. They used 20 different drivers, ten HV drivers and ten PC drivers. First, each car-following period was segmented using segment length, defined as time length of each segment, and segment centroid, defined as the average of the data points within a segment. The centroid has eight dimensions: average longitudinal acceleration, average lateral acceleration, average yaw rate, average vehicle speed, average lane offset, average range, and average range rate. These variables are accounted for in the centroid calculation by converting all variables to a normal Z-scale. Segment lengths were optimized using Metropolis (initialized with segment lengths of 3s) and centroids were identified using MATLAB. Secondly, clusters were developed, defined as regions of similar behavior, using a k-nearest neighbor algorithm in the SAS JMP software. Last, these clusters of behavior were calibrated using the GHR model. A genetic algorithm was applied to minimize the RMSE between the simulated and observed model outputs (velocity). The calibration procedures applied to the clustered data resulted in a much lower RMSE compared to a calibration procedure applied to all drivers, a single driver, and even a single car-following period. Additionally, this research showed significant heterogeneity among PC drivers, where there is a distribution of clusters that describe each individual car driver; in contrast, truck drivers were observed to be much more homogeneous in car-following behavior (Higgs & Abbas, 2015).

Lochrane, Al-Deek, Dailey, and Krause (2015) used IRV data collected from 64 drivers traversing interstate work zone and non-work zone segments along I–95 as part of

the FHWA Living Laboratory. The authors developed and calibrated a new multidimensional psychophysical car-following framework for driver behavior in work and non-work zones. This framework has five different thresholds that are easily observed in data and have physical meaning. The study shows that four different categories of car-following behavior models exist; they can be segmented by traffic condition (congested, uncongested) and by operational condition (work zone, non-work zone). The study found that the car-following model framework thresholds follow different parameter distributions depending on the traffic and operational conditions; these differences in distributions were found to be statistically significant by KS goodness-of-fit tests. This supports the hypothesis that intra-driver heterogeneity exists both between work zones and non-work zones and between congested and uncongested conditions (Lochrane et al., 2015).

Durrani, Lee, and Moah (2016) applied the NGSIM dataset to calibrate Wiedemann 99 car-following model. The authors investigated differences in vehicle following behavior among PCs, HVs, and motorcycles while accounting for lead vehicle type. Each variable (C0–C9) was calibrated via the definition of the variable. The results were validated on a set of validation data by comparing the cumulative distributions of speed and acceleration between the observed and simulated data. The authors found that vehicle following behavior is significantly different among different vehicle classes and vehicle following pairs (e.g. PC-PC, PC-HV). The authors concluded that different driving behavior parameters should be specified for different following vehicle classes and different leader-follower vehicle pairs. Moreover, they found that the variability of parameters for different vehicle pairs should be considered in the formation of distributions of parameters (Durrani, Lee, & Maoh, 2016).

Kurtc and Trieber (2016) used the reconstructed NGSIM dataset to extensively explore the calibration process for the IDM and the full velocity difference model (FVDM).

They used four different objective functions: relative error gap, absolute error gap, mixed error weighted between relative and absolute error of gap, and absolute error of speed. They used three different calibration methods: local (i.e., calibrated model acceleration compared directly to observed acceleration), global (i.e., simulated trajectory of follower compared to empirical trajectory) and platoon (i.e., dynamics of platoon compared to entire empirical dataset). They calibrated the models by applying the interior point algorithm in the MATLAB optimization toolbox. They compared distributions of parameters obtained with the four different measures for each specific model parameter using a two-sample KS test. The variance of absolute gap error was used to explore inter- and intra-driver heterogeneity. They found that global calibration error rates of the car-following models are considerably lower than what can be achieved through macroscopic calibration. The platoon approach is somewhat higher in terms of calibration error, but still acceptable. The ratio between inter-driver and intra-driver variables was between 0.6% and 0.7% for calibration to speed and 4.7% to 5.9% for calibration to gap (Kurtc & Treiber, 2016).

The aerial data sponsored by TNO also enabled some of the first studies into the existence of multi-anticipatory behavior in traffic. Hoogendoorn and Ossen (2006) observed evidence that the strength of the sensitivity of a driver to their second leading vehicle was approximately half that of their sensitivity to the immediate leading vehicle. In some cases, the sensitivity to the second leading vehicle was observed to be higher than the sensitivity to the direct leading vehicle. The leader-follower pair vehicle types were observed to be a critical explanatory factor for the sensitivity of different drivers to multiple leaders; vehicles following a truck were observed to have a weaker reaction to the second leader than vehicles following a passenger car (Hoogendoorn & Ossen, 2006). Additionally, truck drivers were observed to have the strongest reaction to the second leader. A later study, Hoogendorn, Ossen, and Schreuder (2006), explored the existence of

inter-driver heterogeneity with respect to multi-anticipatory behavior. Using vehicle trajectory-level data collected from a helicopter, the authors obtained estimates for the parameter values of the Bexelius, Lenz, and modified Helly car-following models via the maximum likelihood estimation method. Large variations in parameters indicated the existence of inter-driver heterogeneity. Additionally, the authors noted that neither model was observed to be the best in all cases; that is, all models were necessary to describe all of the drivers in the dataset (Hoogendoorn, Ossen, & Schreuder, 2006). Furthermore, a 2007 study by Hoogendoorn, Ossen, and Schreuder found that accounting for multi-anticipatory behavior significantly improves the performance of the model, with the three-leaders models performing the most robustly. Additionally, they determined drivers are responsive to the relative speed differential between themselves and the second and third downstream vehicle (Hoogendoorn, Ossen, & Schreuder, 2007).

2.4. CHAPTER 2 CONCLUSIONS

Chapter 2 summarizes literature related to microscopic trajectory-level data, existing car-following models, and efforts to characterize driver heterogeneity in naturalistic data. The following points are the key takeaways of Chapter 2:

- Trajectory-level data is a promising emerging new data source for driver behavior research. This can be collected aerially or via instrumented vehicles.
- The application of trajectory-level data for car-following calibration is not a new task. However, there is not consensus in the literature regarding the appropriate framework for using trajectory-level data for calibration.

- Many car-following models have been implemented in commercial microsimulation softwares.
- There is unexplained behavioral heterogeneity in trajectory-level data; it has been observed in speed, headway, and acceleration data.
- There have been few attempts to identify explanatory factors for heterogeneity in driving behavior. These attempts have been limited to heterogeneity between acceleration/deceleration behavior, the use of traffic condition (uncongested and congested) and the use of leading vehicle type to explain heterogeneity.

This dissertation explores the degree to which driver specific attributes can be used to characterize inter-driver heterogeneity in car-following behavior and develop a framework to account for this heterogeneity in microsimulation software. Chapter 3 details the data used in this dissertation and the methods applied to process the data and obtain car-following model calibration parameter estimates for future analyses.

Chapter 3: Data Processing

The ultimate goal of this dissertation is to develop a framework to better account for inter-driver heterogeneity attributable to driver specific attributes in microsimulation models. Thus, to evaluate the degree to which heterogeneity can be captured in car-following models, this dissertation uses estimates for car-following model calibration coefficients as a proxy for driver behavior. This chapter details how these calibration parameter estimates were obtained.

Chapter 3 is split into two subsections. Section 3.1 briefly details the data acquisition process; this includes an overview of the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) data collection effort in Section 3.1.1, the SHRP2 Solutions Implementation Assistance Program (IAP) in Section 3.1.2, and the Wyoming Department of Transportation (DOT) IAP dataset in Section 3.1.3. Next, Section 3.2 discusses the procedures applied to obtain estimated car-following model parameter coefficients. First, a brief summary of the radar data post-processing completed by the Virginia Tech Transportation Institute (VTTI) is provided in Section 3.2.1. The procedure for classifying the data into unconstrained and constrained (i.e., car-following) driving states is discussed in Section 3.2.2. Justification for the selection of three car-following models for calibration is provided in Section 3.2.3. Finally, Section 3.2.4 details the car-following model calibration procedure and solution algorithm applied to obtain optimal calibrated parameter coefficients for later analyses.

This chapter resulted in two peer-reviewed conference publications. A paper detailing the development of the radar-vision algorithm was accepted as a special session paper to the 2018 Institute of Electrical and Electronics Engineers Intelligent Transportation Systems Conference in Maui, Hawaii. A paper detailing the developed

calibration procedure and a small case study was accepted as a regular session paper to the same conference. The citations for these two papers are included below:

Hammit, B. E., **James, R. M.**, & Ahmed, M. M. (2018). Radar-Vision Algorithms to Process the Trajectory-Level Driving Data in the SHRP2 Naturalistic Driving Study. *Proceedings of the 2018 Institute of Electrical and Electronics Engineers Intelligent Transportation Systems Conference,* Maui, Hawaii.

Hammit, B. E., **James, R. M.**, & Ahmed, M. M. (2018). A Case for Online Traffic Simulation: Systematic Procedure to Calibrate Car-Following Models Using Vehicle Data. *Proceedings of the 2018 Institute of Electrical and Electronics Engineers Intelligent Transportation Systems Conference*, Maui, Hawaii.

3.1. DATA ACQUISITION

The dataset used in this dissertation was collected as part of the revolutionary SHRP2 NDS data collection effort between 2010 and 2013. Section 3.1.1 provides a brief overview of the SHRP2 NDS data collection effort. Then, Section 3.1.2 introduces the SHRP2 Solutions IAP, which was intended to help deploy insights gleaned from the SHRP2 NDS into practice. Finally, Section 3.1.3 provides an overview of the Wyoming DOT IAP, which was how this dataset was queried from the larger SHRP2 NDS dataset.

3.1.1. The Second Strategic Highway Research Program Naturalistic Driving Study Data Collection Effort

The SHRP2 was a collaborative partnership between the Federal Highway Administration (FHWA), American Association of State Highway and Transportation Officials (AASHTO), and the Transportation Research Board (TRB) (Blatt et al., 2015; US

Department of Transportation Federal Highway Administration, 2018d). This research program was originally authorized as a part of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) in August 2005. As part of SHRP2, there were four product focus areas: safety, infrastructure renewal, reliability, and capacity (Blatt et al., 2015).

The safety component of SHRP2 represents the first large scale study that aimed at preventing the occurrence of collisions, as opposed to previous effort that sought to decrease the severity and improve the survivability of inevitable crashes. The motivation behind the SHRP2 safety data collection effort postulates that by studying diverse drivers' naturalistic behaviors, professionals can better understand the behaviors that cause and avert crashes (Blatt et al., 2015); this knowledge could then be used to improve driver training programs and develop effective countermeasures for crash avoidance. Towards this end, the SHRP2 sponsored a massive naturalistic data collection effort, known as the SHRP2 NDS, which involved instrumenting the personal vehicles of over 3,400 drivers between the ages of 15 and 94 across six data collection locations: Seattle, Washington; Tampa, Florida; Durham, North Carolina; State College, Pennsylvania; Bloomington, Indiana; and Buffalo, New York (Blatt et al., 2015; Virginia Tech Transportation Institute, 2018b). The VTTI was awarded the contract to collect the SHRP2 NDS dataset and are now maintaining the database; freely available information about the dataset can be found on the InSight webpage (Virginia Tech Transportation Institute, 2018b).

The data acquisition system (DAS) collected visual (e.g., forward facing camera) and quantitative (e.g., vehicle position, velocity, and acceleration) trajectory-level data regarding the behavior of the driver for each trip, from engine start to stop, at a frequency of 10Hz (i.e., ten data points collected per second). The components of the DAS are summarized in Table 3.1. Ultimately, the SHRP2 NDS collected 5.4 million trip files while

amassing over 4,300 years of naturalistic driving behavior for analysis (Hankey, McClafferty, & Perez, 2016; Virginia Tech Transportation Institute, 2018b).

Table 3.1 DAS used for SHRP2 NDS Data Collection (Blatt et al., 2015; Virginia Tech Transportation Institute, 2018b)

Instrumentation	Notes
Four video cameras	Front windshield (forward facing), rear windshield (backward facing) and in-cabin cameras (i.e., to capture driver face and hands)
Accelerometers (3 axis)	Lateral, longitudinal, and vertical
Global positioning system (GPS) unit	Latitude, longitude, elevation, time, and velocity
Forward radar	Relative range (i.e., position) and range rate (i.e., (velocity) of up to seven vehicles in front of the instrumented vehicle
Vehicle network data [Vehicle Controller Area Network bus (CAN-BUS)]	Accelerator, brake pedal activation, automatic braking system (ABS), gear position, steering wheel angle, speed, horn, seat belt information, airbag deployment, and other data

The variables of interest for the exploration of acceleration/deceleration behavior proposed in this dissertation include the following:

- Relative velocity | Reported via the front radar unit's longitudinal range rate;
- Relative following distance | Reported via the front radar unit's longitudinal range;
- Subject vehicle acceleration | Reported via vehicle CAN-BUS; and
- Subject vehicle speed | Reported via vehicle CAN-BUS.

It should be noted that all "relative" values occur between the instrumented vehicle, also referred to as the subject vehicle or following vehicle, and their immediate leading vehicle, also denoted as the target vehicle; the relative velocity values are calculated as subject vehicle (following) minus target (lead) vehicle.

In addition to detailed time-series data about driver behavior, the SHRP2 NDS dataset collected extensive data about the driver through self-reporting surveys. This includes, but is not limited to driver demographics questionnaires, risk perception and taking questionnaires, driving knowledge questionnaires, sensation seeking scale inventory, and visual and cognitive tests. The driver surveys and time-series datasets are easily linked through the unique driver ID primary key. For an exhaustive list of the data available about the NDS drivers, see data dictionaries available through InSight (Virginia Tech Transportation Institute, 2018b). For a list of the driver attributes of interest to this dissertation, see Table 3.2.

Table 3.2 Subcategories Comprising the Driver Attributes

Driver Attribute	**Subcategories of the Driver Attribute**
Gender	Male, Female
Age (years)	20–24, 25–29, 30–34, 35–39, 40–44, 45–59, 60–69, 70+
Race	Caucasian, Not Caucasian
Educational Attainment	No College Degree, College Degree, Graduate Degree
Marital Status	Single, Unmarried Partners, Married, Divorced, Widow(er)
Living Status	Live Alone, One Parent Household, Two Parent Household
Work Status	Not Working Outside the Home, Part-Time, Full-Time
Income ($)	Under 39k, 40–49k, 50–69k, 70–99k, 100–149k, 150k+
Household Size	1, 2, 3, 4+
Driver Mileage Last Year	0–5k, 6–9k, 10–12k, 13–15k, 16–19k, 20–24k, 25k+

Moreover, there exists a global information system (GIS) based tool which documents detailed information about roadway facilities near the six SHRP2 data collection sites. The Roadway Information Database (RID) collected elements on more than 25,000 miles of roadway. These elements include, but are not limited to, number, width, and classification (e.g., high occupancy vehicle (HOV), turn, through) of lanes; location, number of approaches, and control type (e.g., all-way stop, signalized, roundabout) of intersections; and guardrail/barrier types (Smadi, 2015). The RID and NDS dataset can be easily linked using GPS coordinates as the primary key. Pairing the NDS and RID datasets with the driver specific questionnaires gives an unprecedented glimpse into the complex interactions of the driver, roadway, and environment during baseline and safety critical (e.g., crash, near crash) events.

In order to protect the privacy of the participants, this dataset is not available freely online. Moreover, requests for data are highly specific to the research question and require both Institutional Review Board (IRB) approval and a data use license (DUL) (Virginia Tech Transportation Institute, 2018a). Although the data collection was federally funded, there is a cost associated with querying a specific dataset to answer a scoped research question; the VTTI can be contacted to obtain a quoted price for the query. However, some queried datasets have been made available freely through the SHRP2 InSight Dataverse (Virginia Tech Transportation Institute, 2018c); these datasets still require IRB approval and a DUL. Many of the datasets available on the Dataverse were originally queried as part of the SHRP2 Solutions IAP, which is described in Section 3.1.2.

3.1.2. Research to Deployment: The Implementation Assistance Program

The IAP was developed to help state and local agencies deploy SHRP2 Solutions (US Department of Transportation Federal Highway Administration, 2018b). Seven rounds of the IAP were offered between 2013 and 2016. To date, over 130 million dollars in technical and financial assistance have been offered to agencies deploying SHRP2 Solutions.

As part of the IAP Round 4 solicitations, eleven states were awarded 1.1 million dollars to develop proof-of-concept prototypes for analyzing driver behavior to understand the factors contributing to highway crashes (US Department of Transportation Federal Highway Administration, 2018a). During Phase 1, state agencies and their research partners obtained a reduced NDS dataset to identify contributing factors for crash and near crash events and suggest possible countermeasures. The eleven states selected for Phase 1 are summarized below:

- Florida | Pedestrian safety at signalized intersections;
- Iowa | Roadway departure crashes;
- Michigan | Roadway geometry and speeding;
- Minnesota | Work zones;
- Nevada | Pedestrian safety;
- New York | Pedestrian safety through high visibility crosswalks;
- North Carolina | Horizontal and vertical curvature;
- Utah | Interchange ramps;
- Washington State | Speeding behavior;
- Washington State | Roadway infrastructure (lighting); and
- Wyoming | Adverse weather conditions (US Department of Transportation Federal Highway Administration, 2018c).

Upon completion of Phase 1, state agencies could apply for up to two additional phases of work. Phase 2 funded additional research on a complete dataset. Phase 3 was awarded for the development and implementation of crash countermeasures.

This dissertation applies a subsample of the Wyoming DOT dataset obtained through Phase 2 of the IAP. Section 3.1.3 describes how the Wyoming DOT IAP dataset was queried.

3.1.3. The Wyoming Department of Transportation Implementation Assistance Program Grant

The Wyoming DOT was awarded an IAP grant to explore how drivers respond to adverse weather and road conditions. The ultimate goal of this IAP was to develop a more realistic variable speed limit (VSL) system and better understand how drivers adjust their behavior to compensate for increased crash risk in adverse weather conditions (Ahmed, 2016; US Department of Transportation Federal Highway Administration, 2018c). As such, the data query for this IAP focused on identifying trips that occurred during adverse weather conditions. The University of Wyoming developed three complementary procedures to automatically flag trips occurring in adverse weather conditions in the larger NDS dataset (Ghasemzadeh, Hammit, Ahmed, & Eldeeb, 2018).

In addition to the adverse weather trips, matching trips occurring in clear weather conditions were obtained for use as a baseline. A clear weather trip matched to an adverse weather trip if it was on the same route, with the same driver, and roughly at the same time of day. For every adverse weather trip queried from the SHRP2 NDS for the Wyoming DOT, two clear weather trips were collected.

Ultimately, a sample of 1284 trips completed by 92 drivers was queried from the SHRP2 NDS as part of the Wyoming DOT IAP. The average trip length was 25.8 minutes,

amounting to approximately 1212 hours of driving data for analysis. These trips were segmented by varying severity of precipitation: clear, fog, very light rain, light rain, moderate rain, heavy rain, and snow.

Previous research has characterized the intra-driver heterogeneity attributable to adverse weather conditions (Hammit, Ghasemzadeh, et al., 2018; Hammit et al., 2019). However, a substantial amount of unexplained heterogeneity was observed in the 665 trips occurring in clear weather conditions; summary statistics for these trips are shown in Table 3.3. Thus, this dissertation seeks to characterize the inter-driver heterogeneity observed in trips traversed on freeways during clear weather conditions obtained through the Wyoming DOT IAP.

Table 3.3 Time-Series Data Summary Statistics

	Average	**Standard Deviation**
Trip Duration [min]	26.1	13.8
Number of Driving States (i.e., car-following or non-car-following)	33.7	19.1
Time Spent in Constrained Driving State (i.e., car-following) [min]	7.32	4.46
Velocity in Constrained Driving State (i.e., car-following) [m/s]	28.0	3.71
Time Gap in Constrained Driving State (i.e., car-following) [s]	1.87	0.61
Following Distance in Constrained Driving State (i.e., car-following) [m]	49.5	13.6
Distance Traveled During a Constrained Driving State (i.e., car-following) [m]	368	198

3.2. CAR-FOLLOWING MODEL CALIBRATION FRAMEWORK

This section details the procedures used to transform the time-series radar and CAN-BUS data into car-following model calibration coefficient estimates, which are used through the remainder of this dissertation to gain insight into differences in driving behavior as a function of driver attributes. Section 3.2.1 briefly summarizes the post-processing completed by VTTI to improve the quality of the radar data. Section 3.2.2 discusses the motivation for the development of the radar-vision algorithm to identify states of constrained driving behavior. Section 3.2.3 justifies the selection of the Wiedemann 99 (W99), Gipps, and Intelligent Driver Model (IDM) car-following models for calibration. Finally, Section 3.2.4 details the calibration procedure used in this dissertation.

3.2.1. Virginia Tech Transportation Institute Radar Data Post-Processing

The VTTI was selected as the original SHRP2 NDS secure data enclave (SDE) and performed significant data post-processing to improve the quality of the radar data. A summary of the post-processing efforts relevant to the variables of interest are listed below:

- Removed ghost targets;
- Identified target vehicle's lane classification relative to subject vehicle;
- Calculated lateral range and range rate variables;
- Smoothed longitudinal range and range rate variables via the application of a cubic spline;
- Created continuous targets via joining of objects;
- Corrected DAS time-lag; and
- Reorganized radar data into a more intuitive format (Gorman, Stowe, & Hankey, 2015).

3.2.2. Extraction of Car-Following Segments with Radar-Vision Algorithms

In Chapter 2, previous research that analyzed trajectory-level data for driving behavior insights was discussed; much previous research in this area was conducted using aerially collected data (e.g., drone, helicopter) and floating cars (i.e., GPS data in planned platoons of vehicles); in contrast, the SHRP2 NDS was collected using instrumented research vehicles (IRVs). One of the distinct differences between IRV and aerially collected trajectories is the spatial and temporal scope. Aerially collected trajectories tend to be limited in scope (e.g., 1000m in the I–80 NGSIM dataset), while IRV trajectories extend the entire duration of a driver's trip (i.e., from engine, or DAS, start to stop). Some of the trajectories in the Wyoming DOT sample of data contain more than an hour of continuous driving data, only some of which is defined as car-following. For the 665 trips in clear weather conditions queried through the Wyoming IAP, the average trip length is 26 minutes, with almost 10 minutes spent in a car-following state. Therefore, an added challenge of working with IRV data is developing an efficient and reliable protocol to identify segments of continuous following behavior. In a previous effort, a radar-vision algorithm was developed to identify the presence of a leading vehicle through the available CAN-BUS and radar data; extended details of this algorithm can be found in Hammit and James (2018b).

The radar-vision algorithm is an iterative smoothing algorithm that uses future and historical radar data to identify the target ID, or lack thereof, that is most likely the leading vehicle. At the termination of this algorithm, the starting and stopping timestep for each driving state is known. A driving state can either be unconstrained, no lead target ID, or constrained, with the target ID of the leading vehicle identified.

Thirty-two hours of driving data were watched for manual verification of the radar-vision algorithm; manual verification of the algorithm confirmed an accuracy rate of 96%.

Moreover, the application of this algorithm improved the continuity of the driving states by reducing the number of distinct driving segments by 79.5% and increasing the average duration of driving segments by 79.2%.

After processing each of the trips through the radar-vision algorithm to identify continuous segments of car-following, a moving average filter with a window size of one second was applied to the radar range rate—i.e., relative velocity—to smooth the resulting measurements. Relative velocity was selected as the smoothing variable based on both a literature search and the observation that it was noisier than the relative distance (i.e., processed range) variable. The window size was identified from a sensitivity analysis, as it provided optimal smoothing while maintaining the original data trends. Figure 3.1 shows the results of applying a moving average filter on processed range rate with a window of one second for a random sample of SHRP2 NDS data. Figure 3.1a shows the difference between the relative velocity before and after the application of the moving average filter. Figure 3.1b and Figure 3.1c illustrate the contrast between the psychophysical car-following plane using the processed range rate (Figure 3.1b) and the processed range rate after the moving average filter was applied (Figure 3.1c). Figure 3.1 illustrates the improved clarity of car-following trends as a result of the application of the moving average filter.

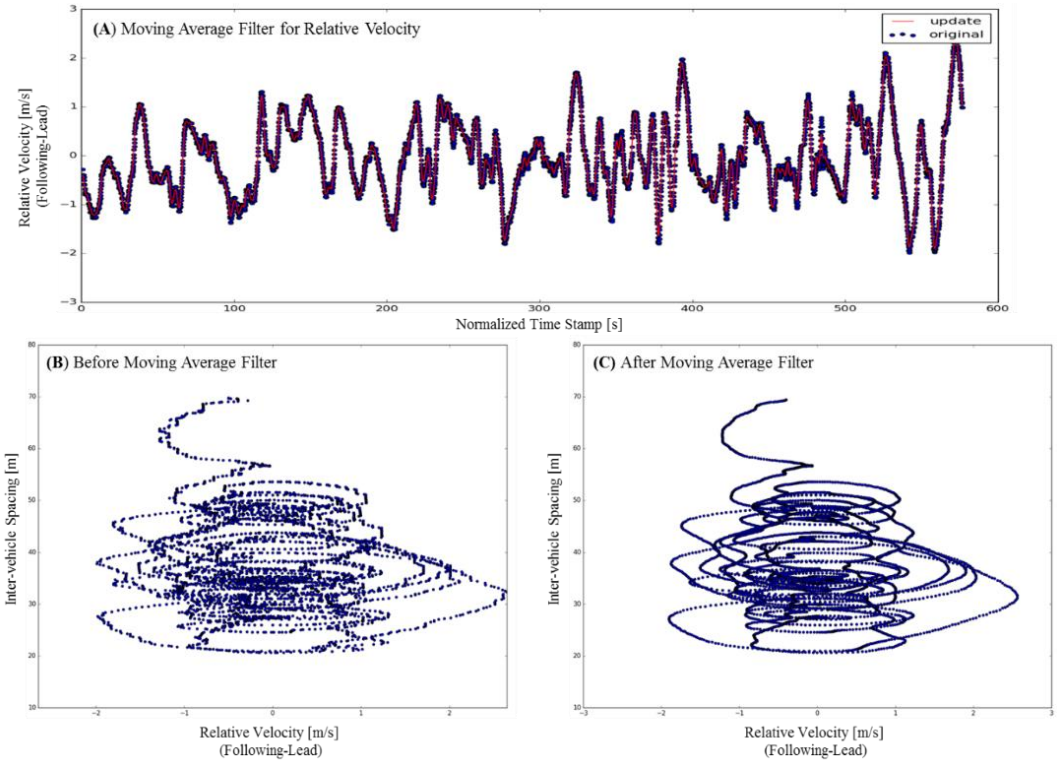


Figure 3.1 Application of moving average filter using relative velocity variable with a one second window (Ahmed et al., 2018)

3.2.3. Car-Following Model Selection

Car-following models are one of the most critical components of microsimulation models because they control the longitudinal, or acceleration, behavior of vehicles; this has a substantial impact on critical model outputs for decision-making, such as capacity. In the literature, a variety of car-following models have been developed and validated (Brackstone & McDonald, 1999; Toledo, 2007). Moreover, multiple efforts have concluded that no one model best fits all observed driving behaviors consistently (Brockfeld et al., 2004; Ossen et al., 2006; Treiber & Kesting, 2013b). Therefore, this dissertation applies three vastly different car-following models in the calibration process:

W99, Gipps, and IDM. These models were selected for their widespread use in practice and because they each represent a different "family" of car-following models. The decision regarding which models get '*to be in the room where it happens*' is documented in this chapter.

3.2.3.1. The Wiedemann 99 Car-Following Model

The W99 car-following model is one of the most frequently used car-following models, given its application in the Planung Transport Verkehr (PTV) Verkehr in Städten – SIMmulationsmodell (VISSIM) software. The W99 car-following model is a psychophysical model; psychophysical models assume that drivers oscillate between regimes of conscious and subconscious reactions to the leading vehicle, depending on the driver's perception thresholds of following distance and relative velocity. A driver's perception is defined as a function of their perceived following distance and speed differential relative to the leading vehicle. Perception thresholds are divided into distinct regimes located on the psychophysical plane (i.e., relative travel speed on the x-axis and following distance on the y-axis) (see Figure 3.2).

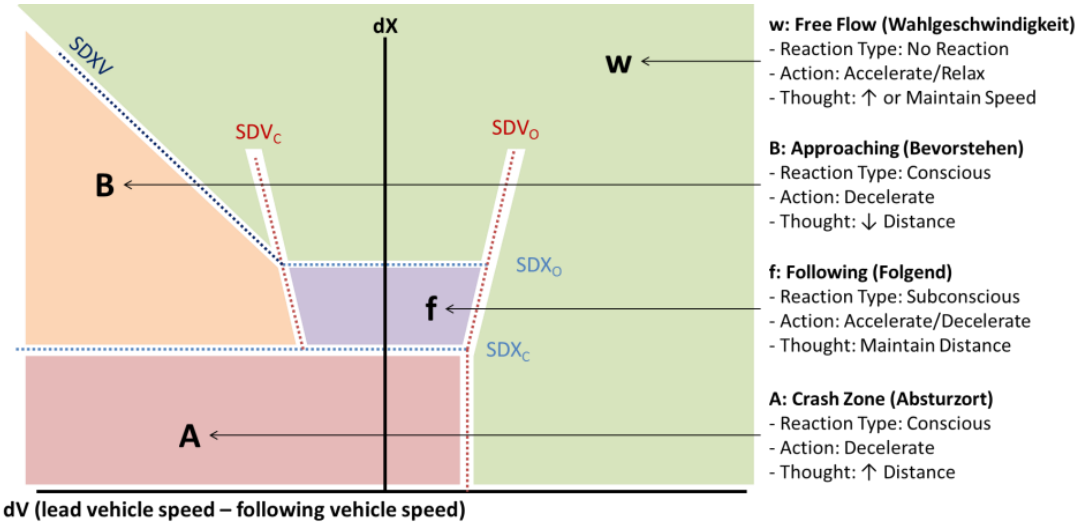


Figure 3.2 W99 Model Regimes on a Psychophysical Plane (Hammit et al., 2019)

Table 3.4 W99 Model Calibration Parameters Interpretations

Variable	**Literature Value** (Liu, 2016)	**Search Range***	**Description**
cc0	1.5	[0.1, 10.0]	standstill distance [m]
cc1	1.3	[0.1, 5.0]	spacing time [s]
cc2	4	[0.1, 15.0]	following variation, max drift [m]
cc3	-12	[-27.0, -5.0]	threshold for entering 'following' [s]
cc4	-0.25	[-5.0, 0.0]	negative following threshold [m/s]
cc5	0.35	[0.0, 5.0]	positive following threshold [m/s]
cc6	0.0006	[0.1, 11.0]	speed dependency of oscillation *6/10000 [10^{-4} rad/s]
cc7	0.25	[0.0, 7.0]	oscillation acceleration [m/s^2]
cc8	2	[0.1, 7.0]	standstill acceleration [m/s^2]
cc9	1.5	[0.1, 8.0]	acceleration at 80kph [m/s^2]
v_des	35	[0.1, 40.0]	desired travel speed [m/s]

*Parameter search spaces were initially set as the bounds recommended in the original model documentation. A random sample of trajectories were selected for sensitivity analysis of the parameter search space. If a majority of the selected trajectories' optimal parameter estimates were on the boundaries of the search space, the search space was expanded; the exception was in cases where expansion of the parameter definition was limited due to physical constraints (e.g., standstill distance should not be negative).

Table 3.5 W99 Model Variable Definitions

Variable	Description
t	Timestamp [s]
T	Time step [s]
seed	random integer to generate unique vehicle behavior [1,1000]
A_F	Predicted following vehicle acceleration for the next timestamp [$A_f = a_f(t+1)$]
Input Data Variables	
$v_L(t)$	Lead vehicle velocity [m/s] at timestamp t
$v_F(t)$	Following vehicle velocity [m/s] at timestamp t
$a_L(t)$	Lead vehicle acceleration [m/s^2] at timestamp t
$a_F(t)$	Following vehicle acceleration [m/s^2] at timestamp t
dV(t)	Relative velocity ($v_L - v_F$) [m/s] at timestamp t
dX(t)	Following distance [m] at timestamp t
Framework Variables	
SDXV(t)	Threshold separating Regime 'B' and Regime 'w' [m] at timestamp t
$SDX_C(t)$	Threshold separating Regime 'A' and Regimes 'B' and 'f' [m] at timestamp t
$SDX_o(t)$	Threshold separating Regime 'f' and Regime 'w' [m] at timestamp t
$SDV_C(t)$	Threshold separating Regime 'B' and Regime 'f' [m/s] at timestamp t
$SDV_o(t)$	Threshold separating Regime 'f' and Regime 'w' [m/s] at timestamp t
Regime Variables	
R(t)	Psychophysical Regime at timestamp t
A	Regime A: Crash/danger zone
B	Regime B: Approaching zone
f	Regime f: Following zone
w	Regime w: Free flow zone

Wiedemann's 1999 car-following model predicts the following vehicle's acceleration and requires eleven input parameters, as shown in Table 3.4. The remaining variables used in the framework (see Figure 3.2), regime (see Figure 3.2), and acceleration calculations are defined in Table 3.5. The framework equations are listed in Equations 3.1 through 3.7:

$$dV(t) = v_L(t) - v_F(t \tag{3.1}$$

$$SDX_c(t) = \begin{Bmatrix} CC_0 & if\ v_L(t) \le 0 \\ CC_0 + CC_1 * v_F(t) & if\ v_L(t) > 0 \wedge (dV(t) \ge 0 \vee a_L(t) < -1) \\ CC_0 + CC_1 * (v_L(t) - dV(t) * (0.5 - b_F)) & if\ v_L(t) > 0 \wedge dV(t) < 0 \wedge a_L(t) \ge -1 \end{Bmatrix} \tag{3.2}$$

$$SDX_o(t) = CC_2 + SDX_c(t) \quad (3.3)$$

$$SDXV(t) = SDX_o(t) + CC_3 * (dV(t) - CC_4) \quad (3.4)$$

$$SDV(t) = CC_6 * dX(t)^2 \quad (3.5)$$

$$SDV_c(t) = \begin{Bmatrix} CC_4 - SDV(t) & if\ v_L(t) > 0 \\ 0 & if\ v_L(t) \leq 0 \end{Bmatrix} \quad (3.6)$$

$$SDV_o(t) = \begin{Bmatrix} CC_5 + SDV(t) & if\ v_F(t) > CC_5 \\ SDV(t) & if\ v_F(t) \leq CC_5 \end{Bmatrix} \quad (3.7)$$

The equation to determine which regime (i.e., free flow, approaching, separating, and crash zone) a driver is in at each timestep is included in Equation 3.8.

$$R(dV(t), dX(t) = \begin{Bmatrix} A & if\ dX(t) \leq SDX_c(t) \wedge dV(t) \leq SDV_o(t) \\ B & if\ dV(t) < SDV_c(t) \wedge dX(t) < SDXV(t) \wedge dX(t) > SDX_c(t) \\ f & if\ dV(t) \leq SDV_o(t) \wedge dV(t) \geq SDV_c(t) \wedge dX(t) \leq SDX_o(t) \wedge dX(t) > SDX_c(t) \\ w & else \end{Bmatrix} \quad (3.8)$$

The acceleration equations for each regime are included in Equations 3.9 through 3.15. Documentation efforts in Liu (2016) and Wiedemann (1996) were used to implement the W99 model for this dissertation.

$$A_F(R(dV(t), dX(t))) = \begin{Bmatrix} A_{F,A} & if\ R(dV(t), dX(t)) = A \\ A_{F,B} & if\ R(dV(t), dX(t)) = B \\ A_{F,f} & if\ R(dV(t), dX(t)) = f \\ A_{F,w} & if\ R(dV(t), dX(t)) = w \end{Bmatrix} \quad (3.9)$$

Regime A

$$A_{F,A} = \begin{Bmatrix} 0 & if\ v_F(t) \leq 0 \\ min\left[a_L(t) + \frac{dV(t)^2}{CC_0 - dX(t)}, a_F(t)\right] & if\ v_F(t) > 0 \wedge dV(t) < 0 \wedge dX(t) > CC_0 \\ min[a_L(t) + 0.5(dV(t) - SDV_o(t)), a_F(t)] & if\ v_F(t) > 0 \wedge dV(t) < 0 \wedge dX(t) \leq CC_0 \end{Bmatrix} \quad (3.10)$$

$$A_{F,A} = \begin{Bmatrix} 0 & if\ A_{F,A} = 0 \\ -CC_7 & if\ A_{F,A} > -CC_7 \\ max[A_{F,A}, -10 + 0.5\sqrt{v_F(t)}] & if\ A_{F,A} \leq -CC_7 \end{Bmatrix} \quad 3.11$$

Regime B

$$A_{F,B} = max\left[\frac{0.5 * dV(t)^2}{SDX_c(t) - dX(t) - 0.1}, -10\right] \quad 3.12$$

Regime f

$$A_{F,f} = \begin{Bmatrix} min[a_F(t), -CC_7] & if\ a_F(t) \leq 0 \\ min\left[max[a_F(t), CC_7], \frac{v_{des} - v_F(t)}{T}\right] & if a_F(t) > 0 \end{Bmatrix} \quad 3.13$$

Regime w

$$A_{F,w} = \begin{Bmatrix} 0 & if\ dX(t) \leq SDX_c(t) \\ CC_7 & if\ dX(t) > SDX_c(t) \wedge R(t-1) = w \\ \min\left[\frac{dV(t)^2}{SDX_o(t) - dX(t)}, CC_8 + CC_9 * min\left[v_F(t), \frac{80 * 1000}{3600}\right]\right] & if\ dX(t) > SDX_c(t) \wedge R(t-1) \neq w \wedge dX(t) < SDX_o(t) \\ CC_8 + CC_9 * min\left[v_F(t), \frac{80 * 1000}{3600}\right] & if\ dX(t) > SDX_c(t) \wedge R(t-1) \neq w \wedge dX(t) \geq SDX_o(t) \end{Bmatrix} \quad 3.14$$

$$A_{F,w} = \begin{Bmatrix} 0 & if\ A_{F,w} = 0 \\ min\left[A_{F,w}, \frac{v_{des} - v_F(t)}{T}\right] & if\ A_{F,w} \neq 0 \end{Bmatrix} \quad 3.15$$

3.2.3.2. The Gipps Car-Following Model

The Gipps car-following model was selected because of its widespread commercial use in the software AIMSUN. The Gipps car-following model is a type of safety distance or collision avoidance model; collision avoidance models are based on Newtonian equations of motion and produce predictions of driver behavior such that the following vehicle can safely react to the leading vehicle should that vehicle decide to come to an abrupt stop.

This dissertation implements the original Gipps model, which was introduced in 1981 and produces driver's desired velocity as an output (Gipps, 1981). This model has six

calibration parameters, all of which have physical interpretations (e.g., desired acceleration, desired deceleration, and minimum following distance); these parameters and their interpretations are listed in Table 3.6. This car-following model predicts the following vehicle's desired velocity subject to two constraints: (i) the desired velocity of the following vehicle (first equation) and (ii) the relative speed and following distance to the leading vehicle (second equation). Table 3.7 summarizes the data inputs required for the Gipps model; these are all easily obtainable from the SHRP2 data. The Gipps equation for desired velocity is shown in Equation 3.16:

$$v_f(t+\tau) = min\left\{\begin{array}{c} v_f(t) + 2.5a_f\tau * \left(1 - \frac{v_f(t)}{V_f}\right) * \sqrt{0.025 + \frac{v_f(t)}{V_f}} \\ b_f\tau + \sqrt{{b_f}^2\tau^2 - b_f * (2\left[x_l(t) - x_f(t) - g_{min}\right] - v_f(t)\tau - \frac{v_l(t)^2}{\widehat{b_l}})} \end{array}\right\} \quad 3.16$$

Table 3.6 Gipps Model Calibration Parameters Interpretations

Variable	**Literature Values** (Gipps, 1981; Olstam & Tapani, 2004)	**Search Range***	**Description**
V_f	35.0	[0.1, 40.0]	desired speed [m/s]
a_f	2.0	[0.1, 4.0]	maximum acceleration [m/s^2]
τ	0.7	[0.1, 2.0]	true reaction time [s]
b_f	-3.0	[-4.0, -0.1]	most severe braking desired (negative value) [m/s^2]
$\hat{b}_l$	-3.5	[-4.0, -0.1]	estimated most severe braking desired of lead vehicle (negative value) [m/s^2]
g_{min}	1.0	[0.1, 10.0]	minimum following distance at stop ($v_f = 0$) [m]

*Parameter search spaces were initially set as the bounds recommended in the original model documentation. A random sample of trajectories were selected for sensitivity analysis of the parameter search space. If a majority of the selected trajectories' optimal parameter estimates were on the boundaries of the search space, the search space was expanded; the exception was in cases where expansion of the parameter definition was limited due to physical constraints (e.g., reaction time should not be negative).

Table 3.7 Gipps Model Variable Definitions

Variable	Description
$v_f(t)$	velocity of following vehicle at time (t) [m/s]
$v_l(t)$	velocity of lead vehicle at time (t) [m/s]
$x_f(t)$	front bumper position of following vehicle at time (t) [m]
$x_l(t)$	back bumper position of lead vehicle at time (t) [m]

3.2.3.3. The Intelligent Driver Model Car-Following Model

The IDM was selected for its prevalent use in practice as the featured car-following model in both the Microscopic Open Traffic Simulation (MOTUS) and Simulation of Urban MObility (SUMO) software. The IDM is a type of social force model; the theory underlying social force models assumes that driving behavior is the resultant of competing forces. The original IDM has two competing forces: (i) the force that encourages the driver to reach and maintain their desired speed and (ii) the force that compels the driver to maintain a desired safe distance from the vehicle it is following.

The IDM predicts following vehicle acceleration and requires six calibration parameters, as shown in Table 3.8; variable definitions used in the model are provided in Table 3.9. This dissertation features the original formulation of the IDM (Treiber, Hennecke, & Helbing, 2000), as documented in Equations 3.17 and 3.18:

$$a_{IDM} = a\left[1 - \left(\frac{v}{v_0}\right)^{\delta} - \left(\frac{s^*}{s}\right)^2\right] \quad 3.17$$

$$s^* = s_0 + \max\left(0, vT + \frac{v\Delta v}{2\sqrt{ab}}\right) \quad 3.18$$

Table 3.8 IDM Calibration Parameters Interpretations

Variable	Literature Values (Kesting, Treiber, & Helbing, 2010)	Search Range*	Description
v_0	35.0	[0.1, 40.0]	Desired speed [m/s]
δ	4	[1, 100]	Free acceleration exponent [unitless]
T	1.5	[0.1, 5.0]	Desired time gap [s]
s_0	2.0	[0.1, 10.0]	Jam distance [m]
a	1.4	[0.1, 4.0]	Maximum acceleration [m/s^2]
b	2.0	[0.1, 4.0]	Desired/comfortable deceleration [m/s^2]

*Parameter search spaces were initially set as the bounds recommended in the original model documentation. A random sample of trajectories were selected for sensitivity analysis of the parameter search space. If a majority of the selected trajectories' optimal parameter estimates were on the boundaries of the search space, the search space was expanded; the exception was in cases where expansion of the parameter definition was limited due to physical constraints (e.g., jam distance should not be negative).

Table 3.9 IDM Variable Definitions

Variable	Description
v	Current velocity [m/s]
Δv	Velocity differential = follower – leader
$\dot{v} = \frac{dv}{dt} = a_{IDM}$	Predicted acceleration for vehicle at time t [m/s^2]
s^*	Desired safe gap [m]

3.2.4. Calibration Procedure

The calibration of car-following models is a complex and resource intensive process. There are two primary schools of thought regarding how to best obtain parameter estimates. The first argues that because car-following models consist of variables that are easily measured in the field (e.g., relative distance, following vehicle speed at time instance *t*, etc.), these parameters should be estimated using the physical values observed in the empirical data; applying the maximum likelihood estimation method to each parameter is one way that this can be achieved (Hoogendoorn & Hoogendoorn, 2010; Hoogendoorn et

al., 2006; Treiber & Kesting, 2013b). Others argue it is best to use an optimization framework to holistically calibrate all of the parameters in a car-following model simultaneously because of the underlying relationships between car-following model parameters (i.e., known correlation between variables) (Ciuffo, Punzo, & Montanino, 2012; Kim & Mahmassani, 2011; Montanino, Ciuffo, & Punzo, 2012; Punzo & Simonelli, 2005). A thorough study by Treiber and Kesting (2013b) provided conclusive evidence that an optimization-based estimation framework is more reliable for car-following model calibration because it is able to capture serial correlations of gaps and speeds. Thus, this research applies the latter method by solving a mathematical model with a nonlinear solution algorithm.

Applying an optimization-based framework for car-following model calibration is no trivial task. As explicitly stated in Ciuffo et al. (2012), the appropriate definition of the optimization problem (i.e., goodness-of-fit function, measure of performance) and identification of solution algorithm is paramount for the successful calibration of car-following models. The measure of performance is the metric used to determine how well the estimated parameters replicate the observed trajectory; temporal headway (Brockfeld et al., 2004; Punzo & Simonelli, 2005), following distance (i.e., relative spacing) (Ciuffo et al., 2012; Punzo, Montanino, & Ciuffo, 2015; Punzo & Simonelli, 2005; Soria, Elefteriadou, & Kondyli, 2014; Treiber & Kesting, 2013b), following vehicle speed (Ciuffo et al., 2012; Punzo et al., 2012, 2015; Punzo & Simonelli, 2005; Ranjitkar et al., 2004; Soria et al., 2014; Treiber & Kesting, 2013b), and following vehicle acceleration (Ossen & Hoogendoorn, 2005; Treiber & Kesting, 2013b) have been used for calibration of trajectories.

Both Treiber and Kesting (2013b) and Punzo and Montanino (2016) provide strong arguments regarding why following distance (i.e., inter-vehicle spacing) is the optimal

choice for measure of performance for this type of problem. Though the use of speed as a measure of performance tends to result in smaller errors, the results are not robust because the error does not adequately propagate through the trajectory. That is, when calibrating to following distance, reasonable following vehicle speeds and acceleration results are produced; the inverse is not true (i.e., negative inter-vehicle spacings are frequently observed when inter-vehicle spacing is not selected as the measure of performance). Thus, inter-vehicle spacing, or following distance, is adopted as the measure of performance for the calibration procedure used in this dissertation. None of the selected car-following models directly produce spacing as an output; therefore, it must be derived from kinematic equations of motion from the model predicted acceleration, for the W99 and IDM car-following models, and velocity, for the Gipps car-following model.

The goodness-of-fit function, otherwise referred to as the objective function, determines the extent to which poor model fit with observed data is penalized. Common selections include root mean square error (RMSE) (Punzo et al., 2012, 2015; Punzo & Simonelli, 2005; Soria et al., 2014), sum of square errors (Treiber & Kesting, 2013b), GEH statistic (Punzo et al., 2012), and Theil's U coefficients (Ossen & Hoogendoorn, 2009; Punzo et al., 2012). An extensive study documented in Ciuffo et al. (2012) confirmed via synthetic data that the RMSE goodness-of-fit function resulted in the best replication of vehicle dynamics, though specifically for the Gipps car-following model. Thus, the objective function for the calibration procedure adopted by this dissertation is shown in Equation 3.19:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(dX_{observed} - dX_{predicted}\right)^2}{N}} \qquad 3.19$$

It is worth noting that this research effort decided to obtain calibration parameter coefficient sets for each collective trip, instead of each individual car-following state (i.e.,

a constrained driving state with an identified leading vehicle), to ensure there was sufficient data completeness to obtain reliable estimates of parameter coefficients, as discussed in Treiber and Kesting (2013b). A complete trip is comprised of two types of driving states: a constrained state with a leading vehicle identified and an unconstrained state. Because of the decision to use inter-vehicle spacing as the measure of performance, this calibration procedure must only consider constrained driving states (i.e., unconstrained driving states have an undefined inter-vehicle spacing). Therefore, the weighted average of the RMSE across all segments of continuous car-following in a trip was reported as the objective function.

It is also worth noting that although only constrained driving states were used for model calibration, there was still sufficient data collected in a free driving state (i.e., without significant influence from a leading vehicle) for parameters such as desired velocity to be properly calibrated. The average maximum following distance of constrained driving states (i.e., car-following) exceeded 90m (more than four car lengths between vehicles), which should conceivably allow drivers to behave without concern of their distant leading vehicle. Moreover, a data completeness framework based on the Wiedemann driving regimes (e.g., free driving, approaching, following, and crash zone) was developed and applied as a method to assess the "completeness" of the dataset; this framework validated that all trips had data in these different driving regimes.

Finally, the solution method determines how the space of feasible parameter coefficients will be searched to identify the set of parameters that best match the observed data in accordance with the selected measure of performance and goodness-of-fit function. There has been a lot of success reported by researchers applying nonlinear solution heuristics. Common solution methodologies include genetic algorithms (Ciuffo et al., 2012; Kesting & Treiber, 2008; Ranjitkar et al., 2004), OptQuest Multistart (Ciuffo et al.,

2012; Punzo et al., 2012, 2015), and Downhill Simplex (Kim & Mahmassani, 2011; Punzo et al., 2012). A literature review revealed that genetic algorithms are the most commonly used heuristic for calibrating car-following models (Ciuffo et al., 2012). It was also documented that although genetic algorithms are not an exact solution algorithm, they are capable of identifying the true values of model parameters in synthetic data (Punzo et al., 2012). Therefore, this dissertation uses a genetic algorithm to calibrate the Gipps, IDM, and W99 car-following model by identifying the sets of model parameters that minimize the RMSE between the predicted and observed inter-vehicle spacing profiles across the car-following states comprising a complete trip. This is summarized in Figure 3.3. Details on the genetic algorithm, which was developed using the Distributed Evolutionary Algorithm in Python (DEAP), are included in Hammit and James (2018a).

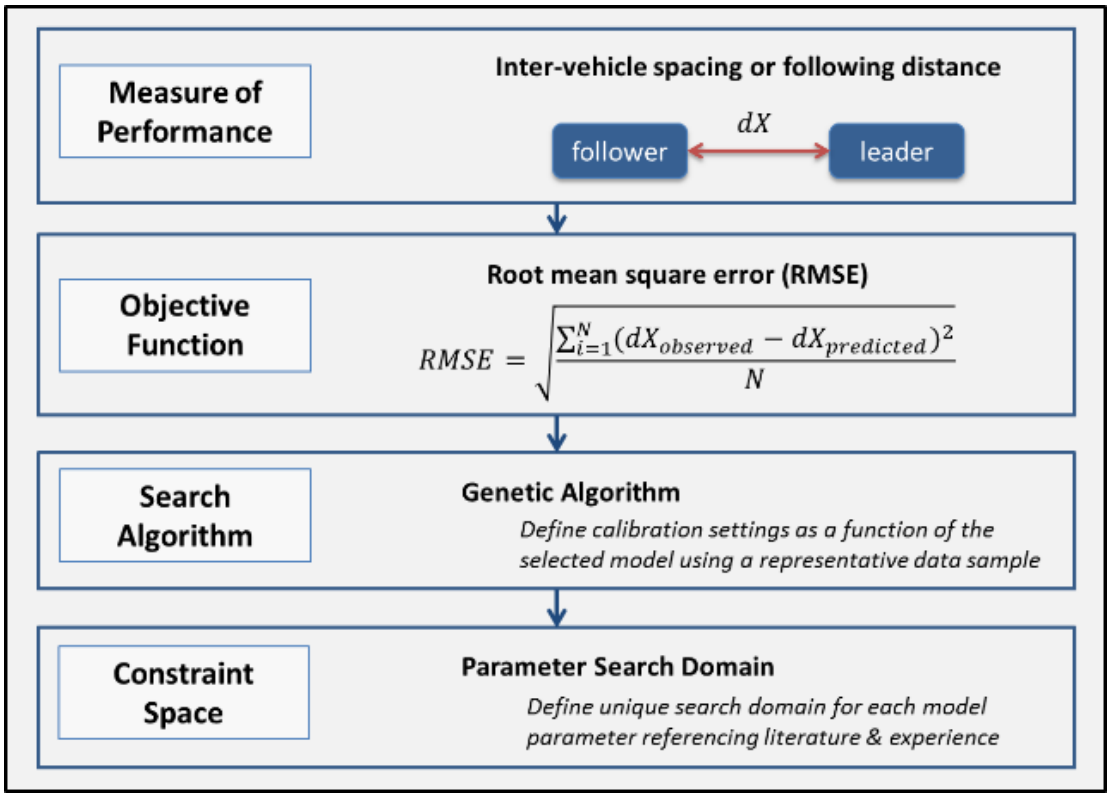


Figure 3.3 Summary of Calibration Procedure (Hammit et al., 2018a)

3.3. CHAPTER 3 CONCLUSIONS

Chapter 3 details the data acquisition and data processing required to transform the time-series radar and CAN-BUS data into the trip-specific optimal calibration coefficient estimates of interest to this dissertation. Section 3.1—Data Acquisition—includes a summary of the SHRP2 NDS data collection effort (Section 3.1.1), the SHRP2 Solutions IAP (Section 3.1.2), and the data queried through the Wyoming DOT IAP grant (Section 3.1.3). Section 3.2—Car-Following Model Calibration Framework—details the radar data post-processing completed by VTTI (Section 3.2.1), the identification of constrained driving states through a radar-vision algorithm (Section 3.2.2), the justification for the

selection of three car-following models for calibration (Section 3.2.3), and the applied calibration procedure (Section 3.2.4).

The key points of Chapter 3 are summarized as follows:

- The data used for this dissertation is a subsample of the SHRP2 NDS dataset. The data was queried through the Wyoming DOT IAP.
- The data used in this dissertation was collected on freeways during clear weather conditions. Six hundred sixty-five trips were identified when the aforementioned filtering criteria were applied to the Wyoming DOT IAP dataset. These trips were completed by 82 drivers. These trips spent, on average, almost 10 minutes in a constrained driving state (i.e., car-following).
- Constrained driving states were extracted using a radar-vision algorithm, an iterative smoothing algorithm that identified periods of homogeneous driving states (e.g., no leading target; leading target ID = 1).
- The W99, Gipps, and IDM car-following models were calibrated as part of this dissertation.
- The calibration procedure used in this dissertation produced a set of estimated car-following model calibration coefficients for each trip in the dataset. The calibration procedure identified the set of car-following model parameter coefficients that minimized the RMSE between the predicted and observed following distance between two vehicles across a trajectory using a genetic algorithm.

Chapter 4: Evidence of Inter-Driver Heterogeneity in the Second Strategic Highway Research Program Naturalistic Driving Study Dataset (Task 1)

This chapter seeks to determine if there is sufficient heterogeneity in the available sample of the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS), described in detail in Chapter 3, to warrant a dissertation dedicated toward understanding and characterizing the unexplained variability in driving behavior. As explained in Chapter 3, this dataset is already controlling for some intra-driver behavioral heterogeneity attributable to road type and weather conditions; this is because the data for this dissertation was queried such that it only contains trips taken on freeways in clear weather conditions. This chapter looks at both trip statistics from the original NDS data sample and the calibrated parameter estimates for the Wiedemann 99 (W99), Gipps, and Intelligent Driver Model (IDM) car-following models.

4.1. Variation Across Trip Statistics

Figure 4.1 illustrates the distributions of trip statistics for the car-following states identified in each trip of the available NDS sample used in this dissertation. The selected statistics include the minimum and maximum acceleration rates and the average acceleration rate, following distance, relative velocity, and time headway. What is immediately evident with a visual inspection of Figure 4.1 is the spread of data points across the distributions. The curves are roughly bell shaped, with some trip statistics more strongly skewed than others. The most symmetric trip statistics are the mean acceleration rate, maximum acceleration rate, and relative velocity, with skewness values around $\pm$ 0.4. The minimum acceleration rate is most strongly skewed left, while the time headway

distribution is the most strongly skewed right. None of the trip statistics report large kurtosis values, which is one indication that this dataset does not have significant outliers.

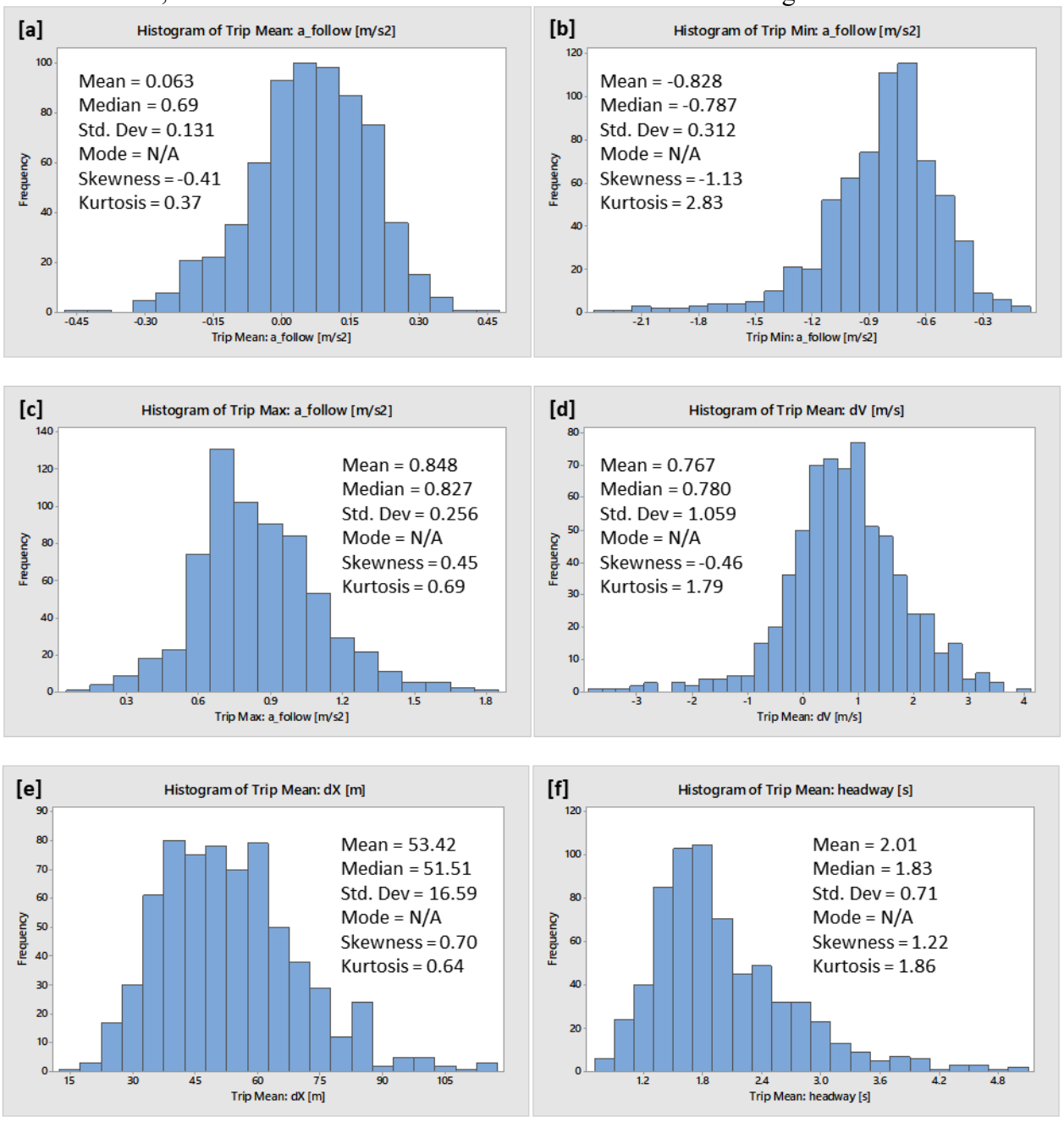


Figure 4.1 Distributions of Trip Statistics

4.2. VARIATION ACROSS INDIVIDUAL PARAMETER ESTIMATES

Figure 4.2 portrays the variation in the estimated W99 calibration coefficients; for additional details on this model, see Section 3.2.3.1. There is substantial visible variation in the data across most distributions of estimated parameter coefficients. The modes of the distributions occurred at one of the boundaries of the search space for the following variation, maximum drift (cc2), threshold for entering following (cc3), negative following threshold (cc4), positive following threshold (cc5), and oscillation acceleration (cc7) distributions (see Figure 4.2). Moreover, the distributions of these estimated parameter coefficients are less bell shaped in nature than the trip statistics. The standstill distance (cc0) and desired velocity (v_des) distribution are the most symmetric, with skewness values below ± 0.5. The most asymmetric distributions are the negative following threshold (cc4) and acceleration at 80 kph distributions. The highest kurtosis values are reported for the spacing time (cc1), negative following distance (cc4), and the acceleration at 80 kph (cc9) distributions. Figure 4.2 does confirm that there is significant spread across the distributions of estimated calibration parameter coefficients. However, Figure 4.2 also illustrates that these distributions are far from normally distributed; this should be accounted for when hypothesis tests are conducted in later chapters of this dissertation (i.e., this violates the assumption of normality required for some statistical tests, such as the one-way analysis of variance).

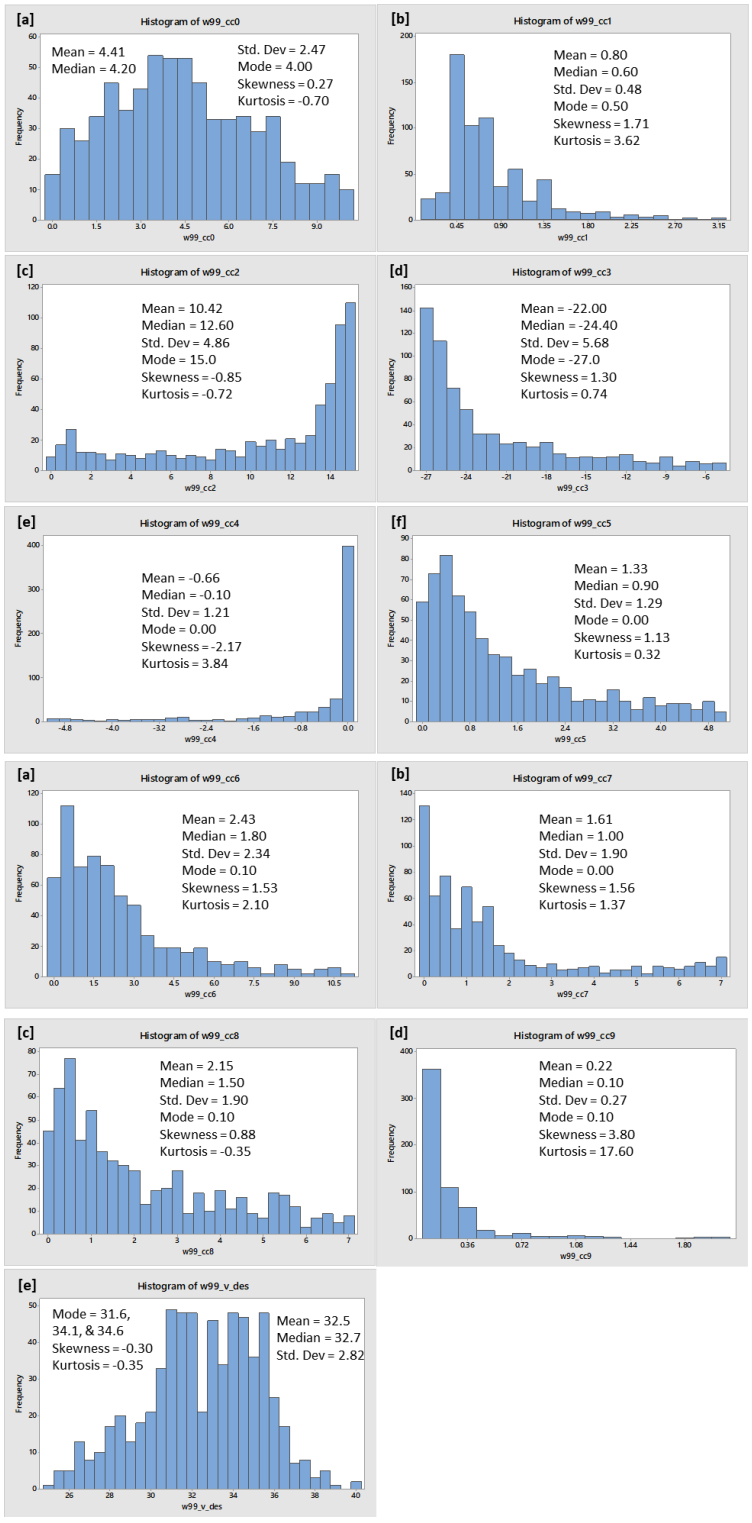


Figure 4.2 Distributions of W99 Car-Following Model Estimated Parameter Coefficients

Figure 4.3 shows the distributions of estimated calibration parameter coefficients for the Gipps car-following model; for additional details on this model, see Section 3.2.3.2. Figure 4.3 confirms that there is substantial variation across the distributions of calibration parameter estimates. These distributions are substantially less skewed than the W99 calibration parameter coefficients, evident by the skewness and kurtosis values. The estimated desired velocity (v_des) coefficient distribution is the most symmetric (skewness value close to 0); this is consistent with observations made about the W99 model. The desired following vehicle deceleration (d_des), perceived desired deceleration of the leading vehicle (d_lead), and minimum following distance at a stop (g_min) parameters were also classified as having fairly symmetric distributions ($-0.5 <$ skewness index < 0.5). The only parameter estimate that appears to be highly skewed is the maximum desired acceleration (a_des) distribution. The kurtosis values are relatively low across all the estimated parameter coefficient distributions, indicating that these distributions are not troubled by outliers (i.e., numerous estimates on the distribution tails). It does not appear that most calibration parameter distributions are normally distributed, which is consistent with what was observed with the W99 model and has implications for the forthcoming hypothesis tests.

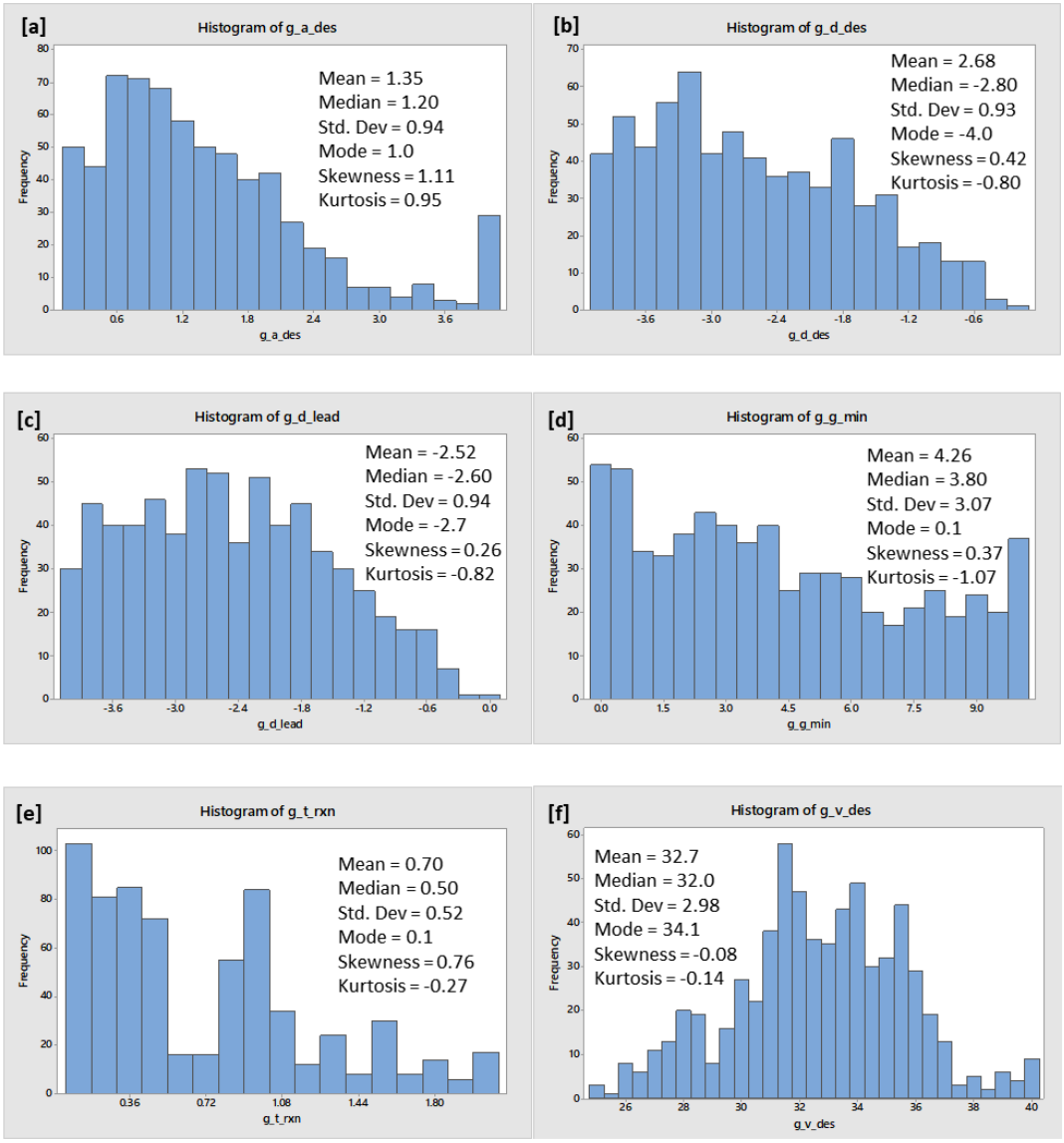


Figure 4.3 Distributions of Gipps Car-Following Model Estimated Parameter Coefficients

Figure 4.4 portrays the distributions of estimated calibration parameter coefficients for the IDM; for additional details on this model, see Section 3.2.4. Visual inspection of Figure 4.4 confirms that there is substantial variation across the distributions of calibrated

parameter estimates. The desired velocity estimated parameter distribution is the most symmetric, with a skewness value of 0.004; this is consistent with observations made about both the W99 and Gipps models. The estimated maximum desired deceleration (b) and the free acceleration exponent (δ) coefficient distributions reported low skewness indices. The distributions of IDM calibration coefficients all have relatively low kurtosis values; this is consistent with the Gipps distributions of calibration coefficients. However, half of the IDM calibration coefficient distributions have modes which fall on the boundaries of the search space (see Figure 4.4): desired deceleration (b), free acceleration component (δ), and desired velocity (v_des); this was similar to the modal observations reported for the W99 model. It does not appear that most parameter distributions are normally distributed, which is consistent with observations made about the W99 and Gipps models.

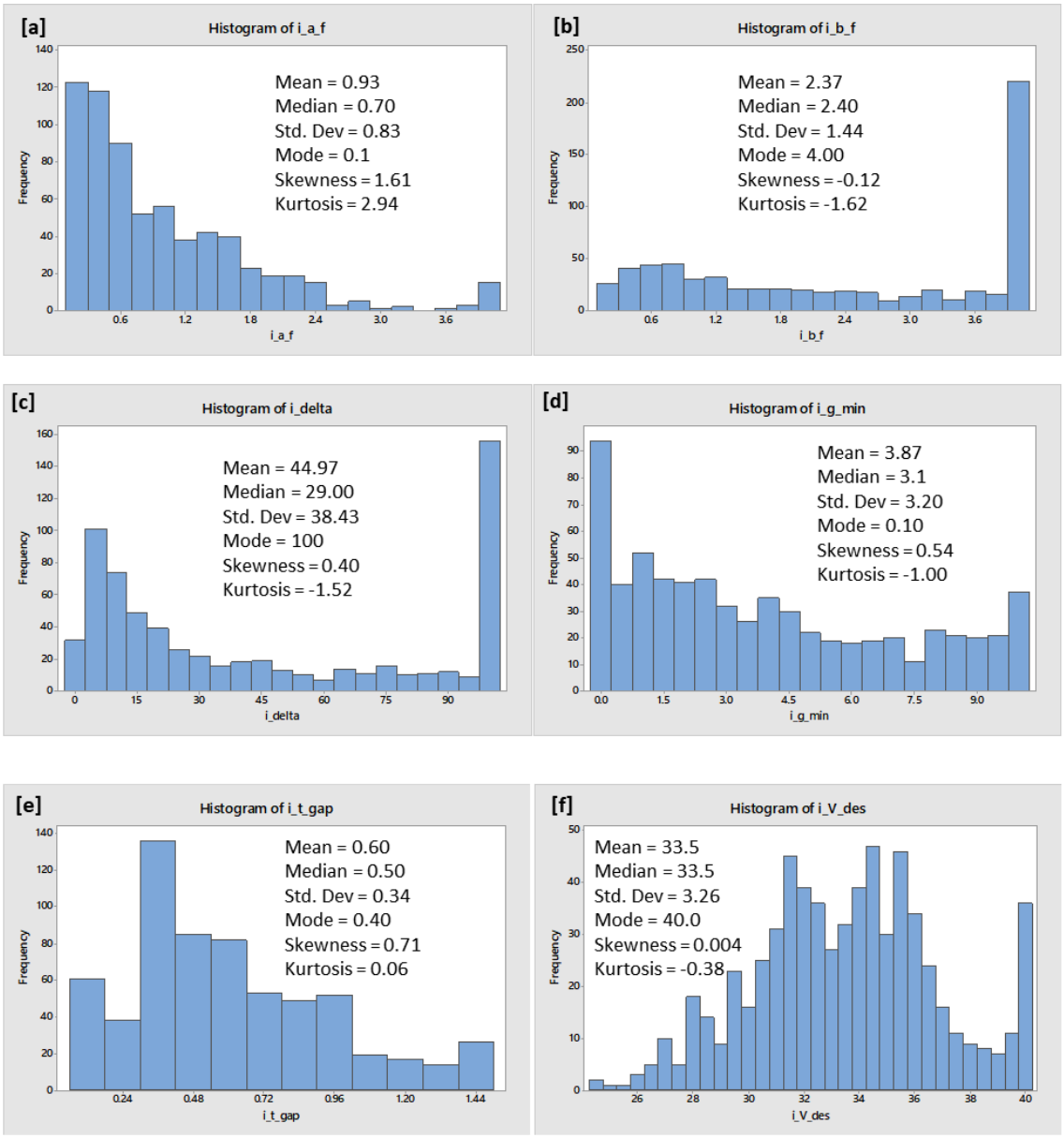


Figure 4.4 Distributions of Intelligent Driver Model Car-Following Parameter Estimates

4.3. VARIATION ACROSS PARAMETERS WITH SIMILAR PHYSICAL INTERPRETATIONS

The three car-following models selected for analysis in this dissertation are comprised of a total of 23 different calibration parameters. Although the three car-following models have different functional forms, several calibration parameters across the models have similar, though not exact, physical interpretations. Therefore, it is an interesting exercise to see how the parameters with similar physical meanings relate across the different model structures. In addition, consistent trends and estimates for these similarly defined parameters helps to increase confidence in the model calibration procedure, which was optimized with a heuristic instead of an exact solution method. There are five physically interpretable parameters that are seen in two or more of the calibrated models:

- Time gap: (i) W99 spacing time (cc1) and (ii) IDM desired time gap (t_gap)
- Maximum acceleration rate: (i) Gipps maximum desired acceleration rate (a_des) and (ii) IDM maximum desired acceleration rate (a);
- Maximum deceleration rate: (i) Gipps maximum desired deceleration rate (d_des) and (ii) IDM maximum desired deceleration rate (b);
- Inter-vehicle spacing at a stop: (i) W99 standstill distance (cc0), (ii) Gipps minimum gap at a stop (g_min), and (iii) IDM jam distance (g_min); and
- Desired velocity: (i) W99 desired velocity (v_des), (ii) Gipps desired velocity (v_des), and (iii) IDM desired velocity (v_des).

Figure 4.5 shows the parameters related to time gap: (i) W99 spacing time (cc1) and (ii) IDM desired time gap (t_gap). The distributions have similar measures of central tendency (i.e., mean, median, and mode). Moreover, both distributions are skewed right, although the estimated W99 spacing time coefficient distribution shows stronger skewedness than the IDM desired time gap coefficient distribution. The estimated IDM

desired time gap coefficient distribution achieves a much smaller kurtosis value than the W99 spacing time distribution, likely attributable to the smaller range of estimated values (i.e., range of estimated W99 spacing time coefficients is [0m, 3.15m), while the range of estimated IDM time gap coefficients is [0m, 1.50m)).

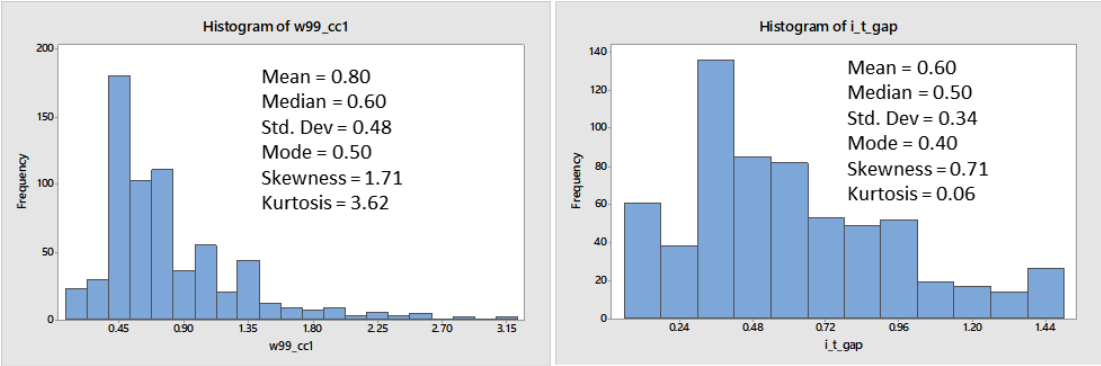


Figure 4.5 Calibration Parameters that are Interpreted to Represent Desired Time Gap

Figure 4.6 compares the Gipps (a_des) and the IDM (a) maximum desired acceleration rate estimated coefficient distributions. There is quite a bit of disparity between the distributions, despite their similar physical interpretations. The median of the Gipps maximum desired acceleration rate coefficient distribution is almost double that of the comparable IDM parameter distribution. However, the models have approximately the same range of estimated values and are both considered strongly skewed to the right. The primary difference appears to be in the distribution mode. The mode of the estimated Gipps maximum desired acceleration rate coefficient (1.0 m/s^2) is an order of magnitude higher than the IDM maximum desired acceleration rate coefficient distribution mode (0.1 m/s^2). Moreover, the Gipps maximum desired acceleration rate coefficient distribution has a higher number of estimates towards the upper bound of the parameter distribution than the IDM maximum desired acceleration rate coefficient distribution. These results are

considered moderately corroborative given the overall similarities in the distribution structures.

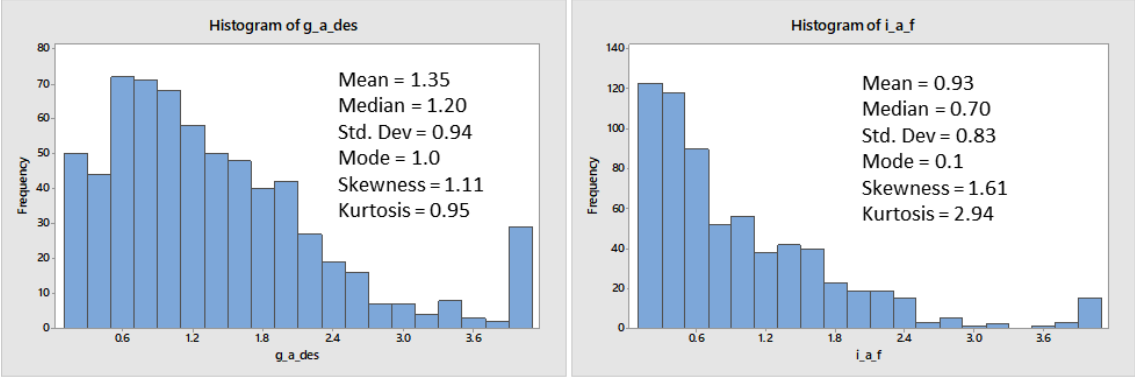


Figure 4.6 Calibration Parameters that are Interpreted to Represent Maximum Desired Acceleration Rate

Figure 4.7 compares the Gipps (d_des) and the IDM (b) maximum desired deceleration rate coefficient estimate distributions. It is important to note that comparisons should be made between the absolute values of the parameter estimates, as the Gipps estimates are negative while the IDM estimates are positive (this is due to differences in the model functional forms). For these distributions, the modes are identical and occur at the parameter search space boundary. The distribution medians are relatively close in value (i.e., 2.80 m/s^2 vs. 2.40 m/s^2). Additionally, both distributions are considered fairly symmetric, with skewness values between ± 0.5. Moreover, both distributions appear to be inversely related to the desired acceleration distribution: that is, the desired acceleration parameter distributions are more skewed towards smaller acceleration rates, while the desired deceleration parameter distributions are more skewed towards larger deceleration rates. However, the IDM distribution for desired deceleration rate has a kurtosis value double that for the Gipps distribution; this is because of the higher number of instances of

maximum desired deceleration rates of 4.0 m/s^2. These differences in parameter coefficient distributions, despite similarities in parameter physical interpretations, could be attributable to differences in the correlation structures of the underlying model parameters. For example, it is well known that the Gipps desired deceleration parameter (d_des) is highly correlated with the Gipps parameter that captures the following vehicle's perceived desired deceleration rate of the leading vehicle (d_lead).

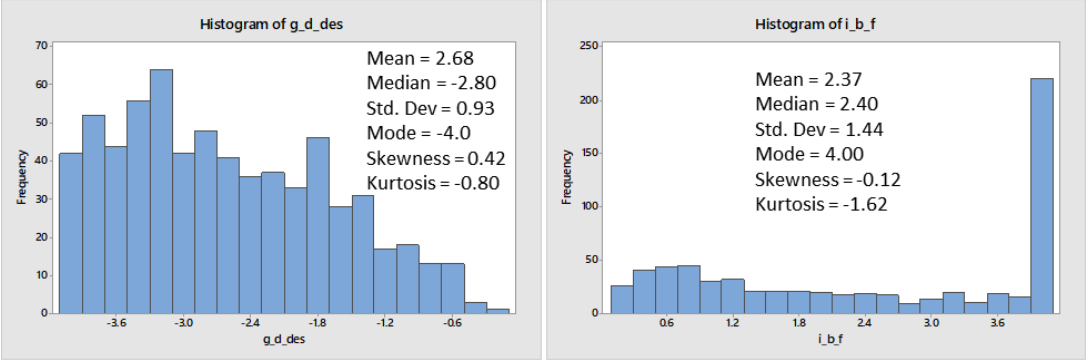


Figure 4.7 Calibration Parameters that are Interpreted to Represent Maximum Desired Deceleration Rate

Figure 4.8 compares the distributions of the parameter coefficients that capture the minimum desired following distance at low speeds: W99 standstill distance (cc0), Gipps minimum gap at a stop (g_min), and IDM jam distance (g_min). The parameter distributions have approximately the same ranges, with estimates occurring between the boundaries of 0 and 10 m. Moreover, all three distributions are fairly symmetric, with slight right leaning skews. The Gipps model achieved the lowest median value (3.1m), while the W99 model was estimated to have the highest median (4.2 m). The modes of the Gipps and IDM parameter estimates are identical and lie toward the parameter search space boundary (0.1 m), while the mode of the W99 model occurs toward the center of the distribution (4.0 m). These distributions appear dissimilar. The differences in the model parameter estimates

may be due to data incompleteness (i.e., low speed conditions were not well represented in this dataset). This dataset was collected on freeways; thus, it is not guaranteed that each trip experienced sufficient congestion during car-following to force the instrumented vehicle to follow a target vehicle at low speeds (i.e., the physical interpretation of this parameter).

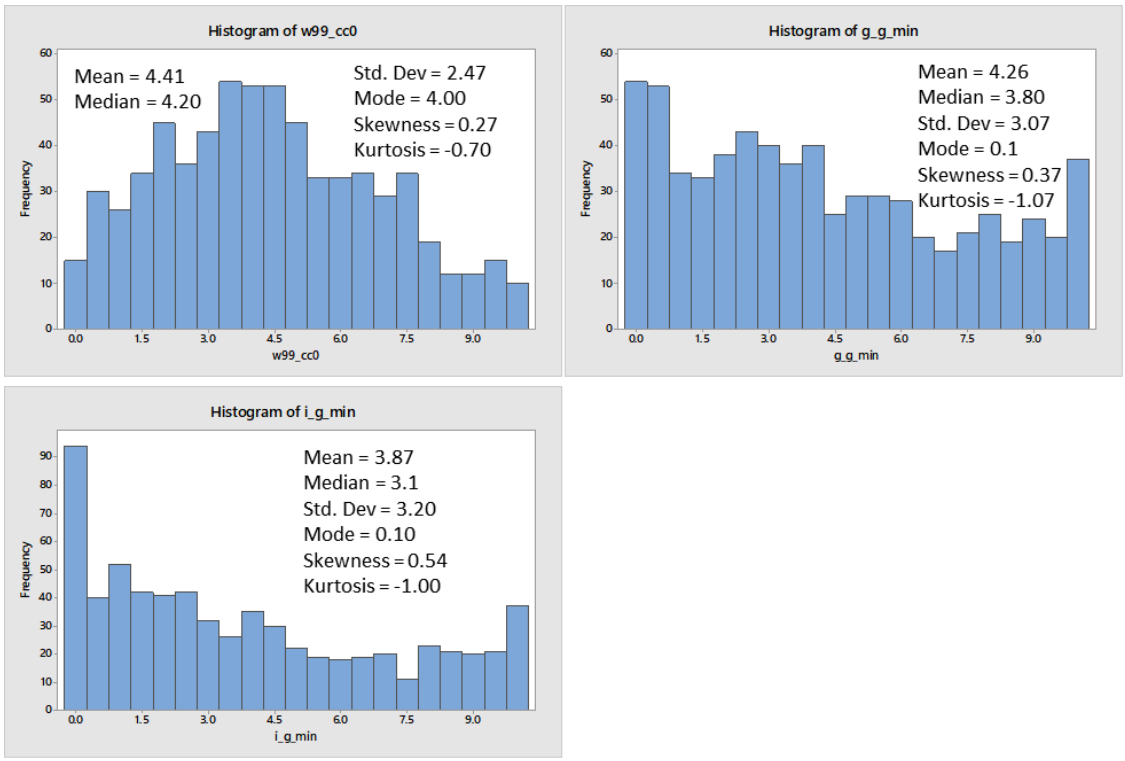


Figure 4.8 Calibration Parameters that are Interpreted to Represent Minimum Inter-Vehicle Spacing

Finally, Figure 4.9 compares the desired velocity parameter distributions for the W99, Gipps, and IDM car-following models. Across all "similarly defined" model calibration parameter definitions, the desired velocity parameter is the only one with an identical definition across models; this has translated to the most similarities in the estimated parameter distributions. The medians across the distributions are markedly

similar, ranging from 32.0 m/s to 33.5 m/s. Moreover, the three parameter distributions are bell shaped in nature, with small skewness and kurtosis indices.

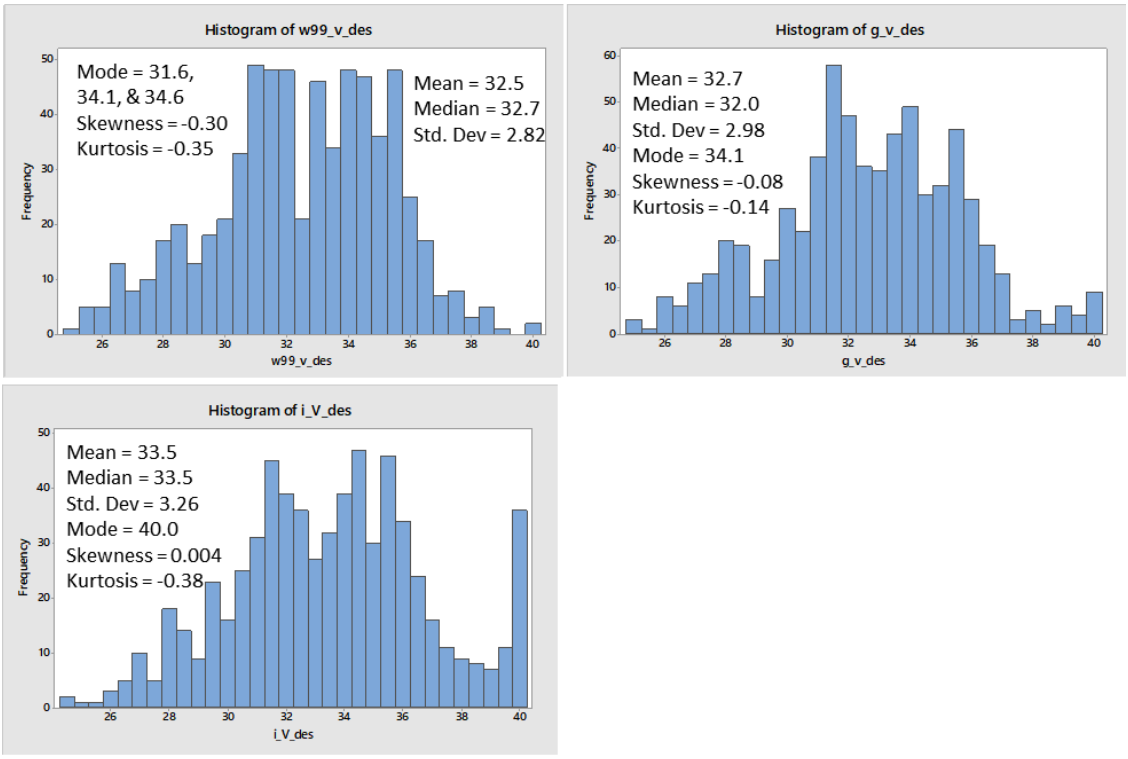


Figure 4.9 Calibration Parameters that are Interpreted to Represent Desired Speed

4.4. CHAPTER 4 CONCLUSIONS

This chapter discusses the variation in distributions of trip statistics and estimated model calibration parameter coefficients in the sample of the SHRP2 NDS dataset available for this dissertation. This dataset was collected in clear weather conditions on freeways, controlling for some known sources of intra-driver behavioral heterogeneity. Key chapter takeaways are summarized below:

- There is significant variation across estimated calibration parameter coefficient distributions. This indicates that there is currently unexplained variation in the dataset that is worthy of further exploration.
- The trip summary statistics are mostly normally distributed.
- The model calibration parameter coefficients are not normally distributed, aside from the desired velocity parameter. This has implications for hypothesis testing completed in subsequent chapters of this dissertation.
- In Chapter 3 it was discussed that the three car-following models used in this dissertation are substantially different in intuition and functional form. However, there are five model calibration parameters that have similar physical interpretations. Aside from the minimum inter-vehicle spacing parameter, which potentially suffers from data incompleteness, these parameters are remarkably similar in measures of central tendency (e.g., mean, median, mode) and distribution shape (e.g., skewness, kurtosis). This provides support in the validity and accuracy of the calibration procedure, which was optimized using a heuristic instead of an exact solution method.

Chapter 4 successfully illustrates that there is unexplained variation of driving behavior in the SHRP2 NDS sample. Chapter 5 begins to explore the degree to which driver specific attributes (e.g., age, gender) can be used to control for the inter-driver behavioral heterogeneity observed in Chapter 4.

Chapter 5: Characterizing the Inter-Driver Heterogeneity Observed in the Second Strategic Highway Research Program Naturalistic Driving Study Dataset (Task 2)[1]

It is illustrated in Chapter 4 that the driving behaviors captured through the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) dataset are diverse, with strong evidence of behavioral heterogeneity. Chapter 5 seeks to determine the degree to which driver specific attributes, such as one's age and gender, help to characterize the observed behavioral heterogeneity. Section 5.1 provides a brief summary of the motivation. Section 5.2 evaluates the differences between different groups of drivers, segmented by driver attributes (e.g., are there statistically significant differences in driving behavior between male and female drivers?). Section 5.3 explores evidence of similarities in driving behavior within a group of drivers segmented by driver attributes (e.g., are there similarities in driving behavior within a group of younger drivers? Moreover, are there notable differences between driving behavior of younger drivers and older drivers?). The analysis conducted to support Section 5.2 was accepted as a peer reviewed special session paper at the 2018 Institute of Electrical and Electronics Engineers Intelligent Transportation Systems Conference in Maui, Hawaii. The analysis conducted to support Section 5.3 was accepted for presentation at the 2019 Annual Meeting of the Transportation Research Board and recommended for publication in the Transportation Research Record. Lastly, a presentation on the development of the homogeneous driver groups was made at

[1] Chapter 5.2 will be published in an upcoming edition of the Transportation Research Record. The full citation is as follows: James, R. M. & Hammit, B. E. (2019). Identifying Contributory Factors to Heterogeneity in Driving Behavior: A Clustering and Classification Approach. *In Press: Transportation Research Record.* B. E. Hammit contributed significantly to the optimization procedure used to obtain calibrated car-following model parameter values. These parameters were used as surrogates for driver behavior in the paper. The application of the clustering and classification algorithms and the interpretation of the model results were completed by R. M. James.

the Seventh International Symposium on Naturalistic Driving Research. The citations for each of these are provided below:

James, R. M. & Hammit, B. E. (2018). Exploring the Use of Driver Attributes to Characterize Heterogeneity in Naturalistic Driving Behavior. *Proceedings of the 2018 Institute of Electrical and Electronics Engineers Intelligent Transportation Systems Conference, Maui, Hawaii*.

James, R. M. & Hammit, B. E. (2019). Identifying Contributory Factors to Heterogeneity in Driving Behavior: A Clustering and Classification Approach. *In Press: Transportation Research Record*.

James, R. M. & Hammit, B. E. (2018). The Role of Driver Attributes in Understanding Driver Behavior. *Seventh International Symposium on Naturalistic Driving Research*. Blacksburg, VA. Podium Session.

5.1. MOTIVATION

Heterogeneity in driving behavior is becoming increasingly well-documented, especially as high-resolution data is becoming more ubiquitously collected. This behavioral heterogeneity can take one of two forms: inter-driver heterogeneity, where a sample of drivers exhibit different behaviors despite similarity in driving conditions (e.g., cautious drivers vs. aggressive drivers), and intra-driver heterogeneity, where the same driver behaves differently as a function of variations in their driving conditions (e.g., one's driving behavior in clear weather conditions vs. their behavior in adverse weather conditions, such as snow). One area in transportation where this phenomenon has been evident is in crash rate analysis. In this sector, it is apparent that one's gender and age influences their likelihood of being involved in a crash. Using crash rates per 100-thousand

drivers, one can see that men have a higher crash rate than women across all crash types; however, male drivers make up a much higher proportion of drivers involved in fatal crashes (75% / 25% split) compared to property damage only (58% / 42% split) (Insurance Information Institute, 2018). Moreover, age also appears to be correlated with crash risk. When normalizing by vehicle miles traveled, it was observed that drivers under the age of 18 and over the age of 80 are the most at-risk age groups for all crash types (Tefft, 2017). Although these statistics do not control for exposure rates, they do indirectly support the hypotheses of inter-driver behavioral heterogeneity.

This dissertation evaluated the heterogeneity in a sample of the SHRP2 NDS dataset obtained through the Wyoming Department of Transportation (DOT) Implementation Assistance Program (IAP) in Chapter 4. The distribution of optimal car-following model calibration coefficients illustrated that there is significant heterogeneity in the data, despite controlling for weather condition and facility type, and that standard car-following models used in microsimulation packages can capture this heterogeneity when properly calibrated. Additionally, as illustrated in Section 2.3.1, there is considerable evidence that inter-driver heterogeneity as a function of driver specific attributes exists in empirically collected data. Previous calibration studies have concluded that the existence of driving behavior heterogeneity constitutes a baseline that no model ignoring the presence of this heterogeneity in the underlying data can outperform (Treiber & Kesting, 2013a).

Presently, state-of-practice calibration efforts account for heterogeneity through the random sampling of driver parameters from known distributions. This dissertation hypothesizes that by controlling for driver attributes we can further segment the driving population into groups exhibiting more homogeneous driving behavior, which are more appropriate to model with a set of representative calibration parameter estimates. Towards

this end, Chapter 5 seeks to characterize the heterogeneity in driving behavior that may be attributable to driver specific attributes, such as age and gender.

There are two primary research questions in Chapter 5. The first seeks to understand how **different** the driving behaviors are **between** different groups of drivers (e.g., do male drivers exhibit statistically significant differences in driving behaviors from female drivers). The second research question seeks to understand how **similar** the driving behaviors are **within** different groups of drivers (e.g., should female drivers be considered a group of homogeneous drivers? Or are their behaviors too inconsistent to be clustered together?). By the conclusion of Chapter 5, evidence will be provided that the larger driving population can be divided into homogeneous driver groups, segmented by different combinations of driver attributes; this will be used for application in later chapters of this dissertation.

5.2. Differences in Driving Behavior Between Subgroups of Drivers Segmented by Driver Specific Attributes

The objective of this chapter is to better understand how **different** the driving behaviors are **between** different groups of drivers (e.g., do male drivers exhibit statistically significant differences in driving behaviors from female drivers). Chapter 5.2 is intended to answer research question 4 of this dissertation. More specifically, this dissertation hopes to understand the following: when subsamples of drivers are created using driver attributes for segmentation, are the parameter sets for the different subsamples sufficiently different to justify the separation of the data (e.g., is the calibrated reaction time of male drivers statistically different than female drivers)?

5.2.1. Methodology

The methodology used to conduct this analysis has three components. The first effort identifies which car-following models to include in the analysis; for details on this effort, the reader is referred to Section 3.2.3. The second effort develops and implements a calibration procedure to identify optimal parameter sets for each trip; for details on this effort, the reader is referred to Section 3.2.4. Next, the calibrated parameter coefficient estimates are split into "subcategories" (e.g., male, female) that comprise a driver "attribute" (e.g., gender). The driver attributes and their comprising subcategories available for the SHRP2 NDS drivers are shown in Table 3.2. Finally, hypothesis tests are conducted to evaluate if there are significant differences between the subcategories that comprise a driver attribute; the selected statistical tests are described next

Kruskal-Wallis one-way analysis of variance (ANOVA) tests were applied to provide insight regarding the statistical differences in driving behavior across drivers categorized by their unique driver attributes; this is a non-parametric alternative to the traditional one-way ANOVA test and is required given the non-normality of the distributions of calibration parameter coefficient estimates, confirmed using Anderson-Darling statistical tests. In this context, the null hypothesis assesses if the medians of the subcategories' calibration coefficients (e.g., reaction time) are equal (i.e., there is not a statistical difference in reaction times between male and female drivers). Thus, small p-values, 5% alpha level selected for this analysis, result in a rejection of the null hypothesis and the conclusion that there are behavioral differences between the subgroups of drivers. It is important to note that the ANOVA p-values do not necessarily indicate a statistically significant difference between every subgroup of drivers, but rather that there is a statistical difference between all of the subgroups of drivers.

This section evaluates the variation in estimated calibration parameter coefficients across all subcategories (e.g., male, female) comprising a driver specific attribute (e.g., gender). There are eight driver specific attributes that are analyzed in this chapter (see Table 3.2): gender, age, race, educational attainment, marital status, income, household size, and driver mileage last year. These driver specific attributes have a variable number of subcategories ranging from two (e.g., gender, race) to eight (e.g., age). The trends in estimated parameter coefficients across subcategories of driver attributes are evaluated from a qualitative and quantitative perspective. The trends are evaluated qualitatively in Figure 5.1 through Figure 5.29 and interesting observations are annotated in the text of this chapter. Quantitative evaluations of estimated parameter coefficients are supported by Table A.1 through Table A.29, included in Appendix A of this dissertation, and are discussed in this chapter.

5.2.2. Results

The results of this subsection are divided into two sections. Section 5.2.2.1 explores trends across important trip summary statistics as a function of driver specific attributes. Section 5.2.2.2 explores trends in estimated parameter coefficients across driver attributes for each individual calibration parameter of the car-following models. Although three different car-following models were used for calibration, some of the parameters are "shared" across models (i.e., they have similar physical interpretations. In Section 5.2.2.3, trends in the shared parameters are described.

5.2.2.1. Trip Descriptive Statistics Results

Figure 5.1 shows how the average acceleration rate of all car-following segments comprising a complete trip varies according to driver attributes. The differences in average acceleration rate vary at a statistically significant level for each of the subcategories that divide a specific driver attribute. This strongly suggests that acceleration behavior in the naturalistic driving data, the output of car-following models, varies between different groups of drivers. However, there are not many conclusive visual trends in the data, particularly for the attributes that are segmented into multiple categories. According to Figure 5.1a, male drivers tend to have smaller average acceleration rates than female drivers; however, the large confidence intervals associated with the male and female subcategories indicate variability within the subcategories. Figure 5.1d suggests that as one's level of education increases, their average acceleration rates decrease. In Figure 5.1e, drivers that identify as single or unmarried partners have approximately equal average acceleration rate, while married, divorced, and widow(er) subcategories of marital status have approximately equal average acceleration rates; the former subcategories exhibit a smaller average acceleration rate than the latter. Finally, there are only two subcategories of data where the average acceleration rate across all drivers belonging to said subcategory are negative: drivers with a graduate degree and drivers with reported incomes exceeding $150k.

The highest within subcategory variation is observed for the widow(er) subcategory of the marital status attribute (standard deviation = 0.21 m/s^2), while the lowest within subcategory variation is observed for the unmarried partners subcategory (standard deviation = 0.03 m/s^2); this is illustrated in Figure 5.1 and documented in Table A.1. The standard deviations of the subcategories that comprise a driver attribute are averaged to assess the variation of parameter estimates across the subcategories (i.e., the within

attribute variation). The attribute variation is remarkably similar across the eight attributes, ranging from 0.12 to 0.13 m/s^2.

Figure 5.2 shows how the average minimum acceleration, or deceleration, rate of all car-following events comprising a complete trip varies according to driver attributes. The race, education, and household size attributes were not found to be statistically significantly different between the respective subcategories. As illustrated in Figure 5.2a, the average female driver minimum acceleration rate is stronger than that observed in the male data. Moreover, Figure 5.2b suggests that as age increases the average minimum acceleration rate decreases (i.e., age and average hard braking value are inversely related).

Table A.2 indicates that the strongest average minimum acceleration rate was observed for the youngest age group (mean = -0.99 m/s^2); this could be related to the inexperience of younger drivers. The least severe average minimum acceleration rate was observed for the widow(er) subcategory (mean = -0.58 m/s^2); the second least severe braking behavior was observed for the oldest age category (mean = -0.64 m/s^2). This could be indicative of risk aversion in driving behavior, possibly attributable to self-regulated risk-taking behavior due to known cognitive and motor skills decline. The most variation within a subcategory of drivers was observed for the under $39k income group (standard deviation = 0.39 m/s^2), while the least variation was observed for the unmarried partner subcategory (standard deviation = 0.21 m/s^2). The within attribute variation was remarkably similar across the different attributes, with average standard deviations ranging from 0.30 m/s^2 to 0.32 m/s^2.

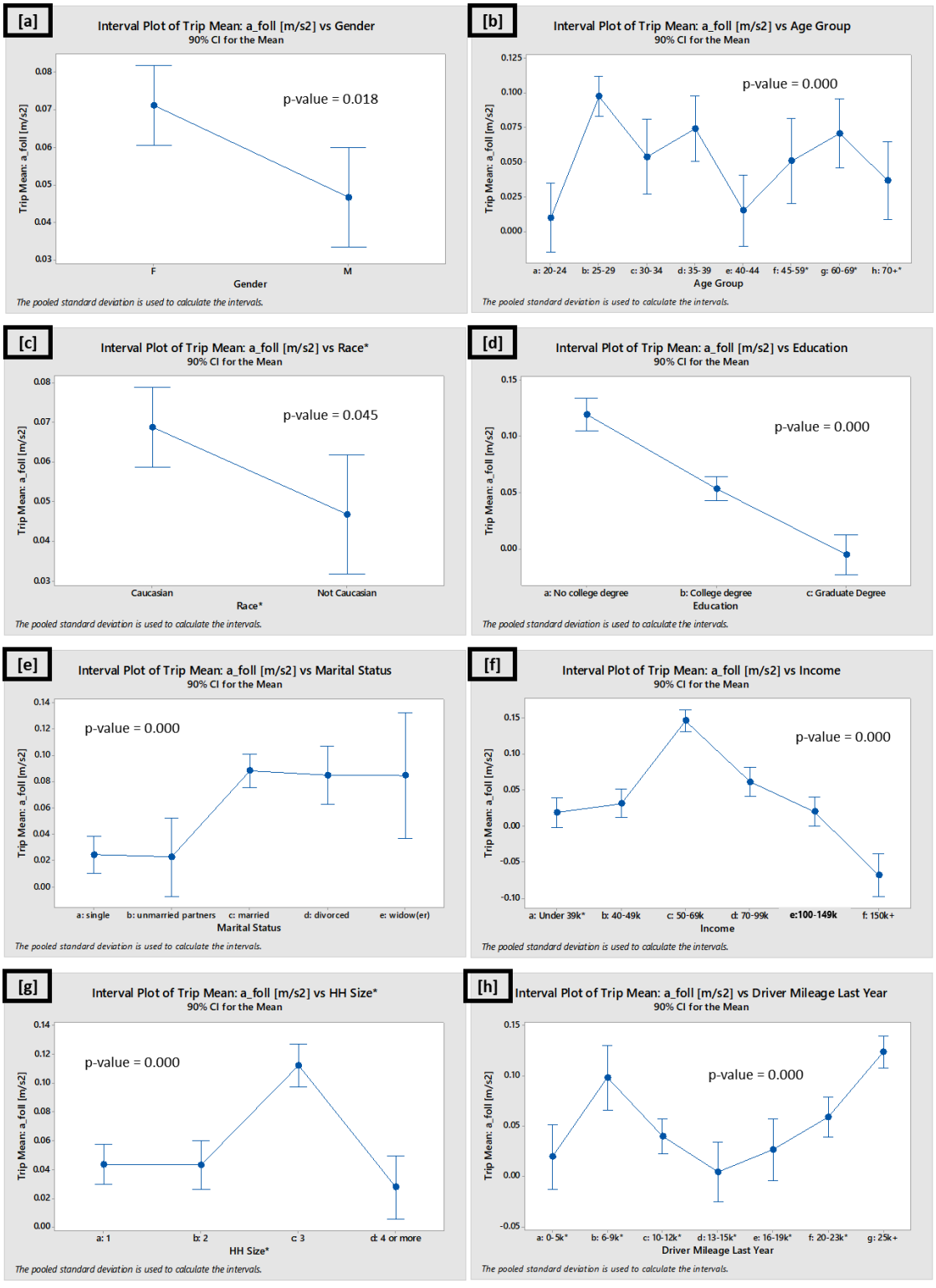


Figure 5.1 Mean Acceleration Segmented by Driver Attributes

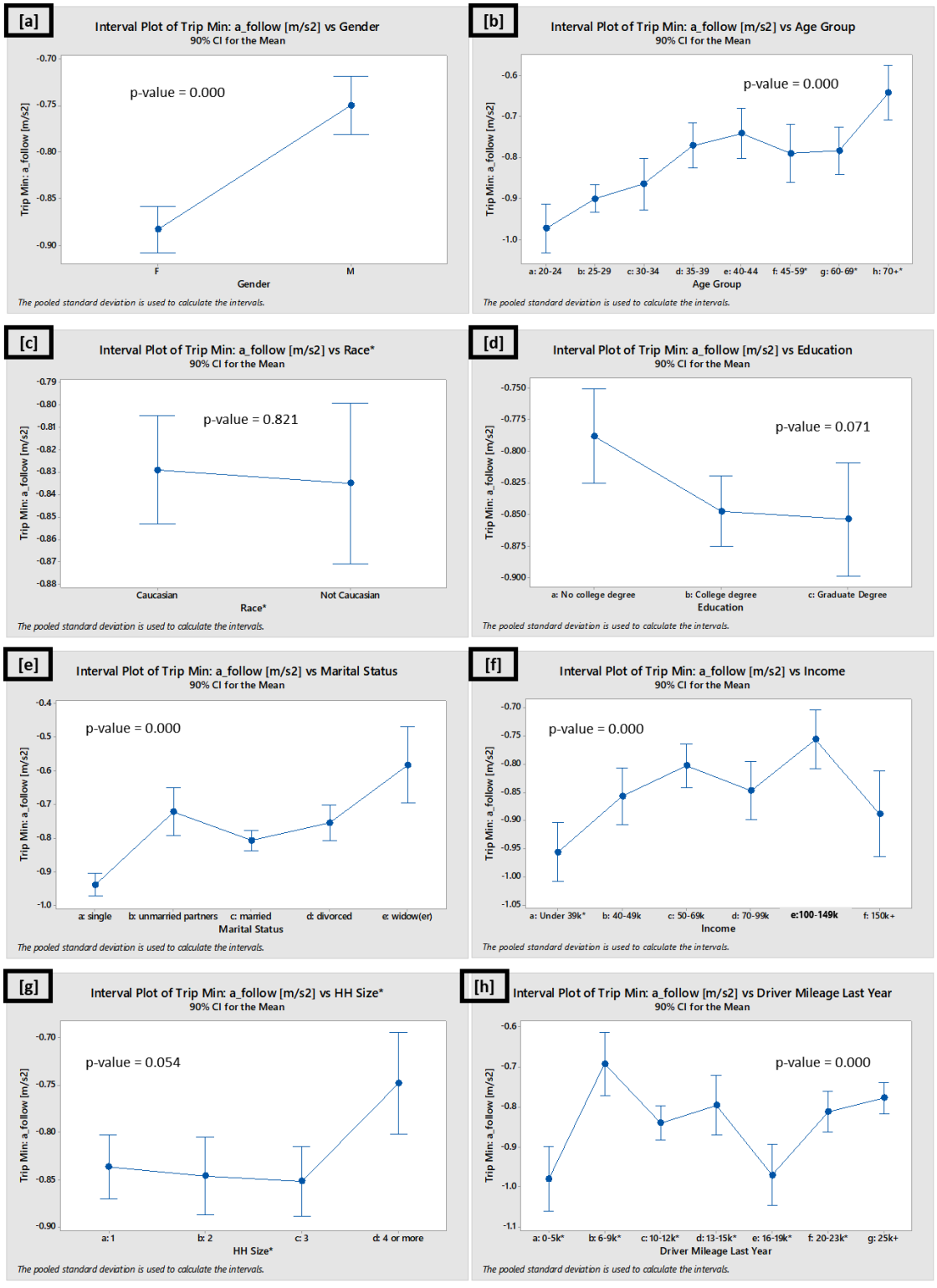


Figure 5.2 Minimum Acceleration Segmented by Driver Attributes

Figure 5.3 illustrates how the average maximum acceleration rate of all car-following events comprising a complete trip varies according to driver attributes. There are statistically significant differences in all subcategories of driver attributes except for the race attribute. Figure 5.3a, suggests that female drivers exhibit a higher average acceleration rate than male drivers. Considering the trends of Figure 5.1a, Figure 5.2a, and Figure 5.3a, it appears that female drivers in this particular NDS sample exhibit stronger acceleration and deceleration rates than male drivers, on average; this could be evidence that female drivers are more aggressive or drive more distracted (e.g., children in vehicle) than male drivers. Figure 5.3b indicates there is a general trend that as age increases the average strongest exhibited acceleration rate decreases. Moreover, as income increases, the average maximum acceleration rate decreases, with exception of the $50–69k subcategory (Figure 5.3f).

As documented in Table A.3, the strongest average acceleration rate was observed with the 25–29-year-old driver subcategory (mean = 0.97 m/s^2), while the smallest average maximum acceleration rate was estimated for the oldest age group (mean = 0.65 m/s^2). The smallest within subcategory variation was observed for the unmarried partners subcategory (standard deviation = 0.12 m/s^2), while the largest within subcategory variation was calculated for the widow(er) subcategory (standard deviation = 0.27 m/s^2). The smallest attribute variation was calculated for the marital status attribute [mean(standard deviation) = 0.22 m/s^2], while the largest attribute variation was observed within the driver mileage last year attribute [mean(standard deviation) = 0.27 m/s^2].

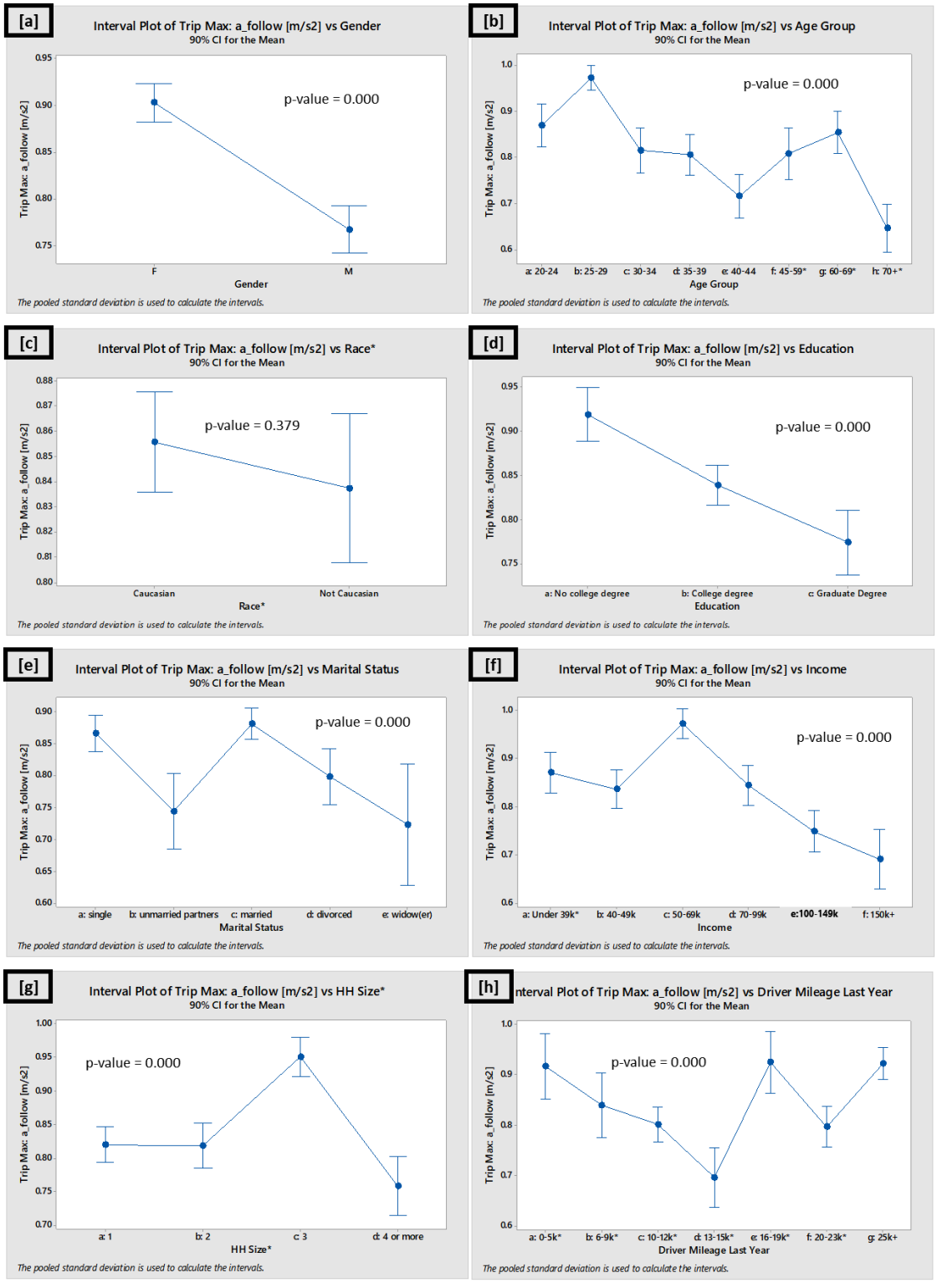


Figure 5.3 Maximum Acceleration Segmented by Driver Attributes

Figure 5.4 portrays how the average relative velocity of all car-following events comprising a complete trip varies according to driver attributes. The relative velocity was calculated as following vehicle speed minus leading vehicle speed; thus, positive values indicate the following vehicle was traveling faster than the leading vehicle. All differences between subcategories of driver attributes are statistically significantly different. As illustrated in Figure 5.4a, male drivers exhibit higher relative velocities than female drivers. Figure 5.4d suggests that as educational attainment increases, relative velocity increases. Moreover, household size and relative velocity appear to be positively correlated (Figure 5.4g).

As documented in Table A.4, the unmarried partners subcategory of marital status has the highest average velocity differential, by a significant margin (mean = 2.2 m/s^2); it was one of two subcategories with mean relative velocities above 1.5 m/s^2, indicating the following vehicle was traveling much faster than the leading vehicle. The widow(er) subcategory reported the smallest average velocity differential (mean = -1.5 m/s^2); it is one of three subcategories to report a negative average relative velocity value, where the leading vehicle was consistently moving faster than the following vehicle (i.e., widow(er), age group 70+, and 6–9k subcategory of driver mileage last year). The 30–34 age group reported the smallest within subcategory variation (standard deviation = 0.5 m/s^2), while the 70+ age group reported the largest variation (standard deviation = 1.7 m/s^2). The marital status attribute had the lowest average standard deviation across all subcategories (i.e., within attribute variation) [mean(standard deviation) = 0.95 m/s^2], while the driver mileage last year attribute reported the highest variation [mean(standard deviation) = 1.07 m/s^2].

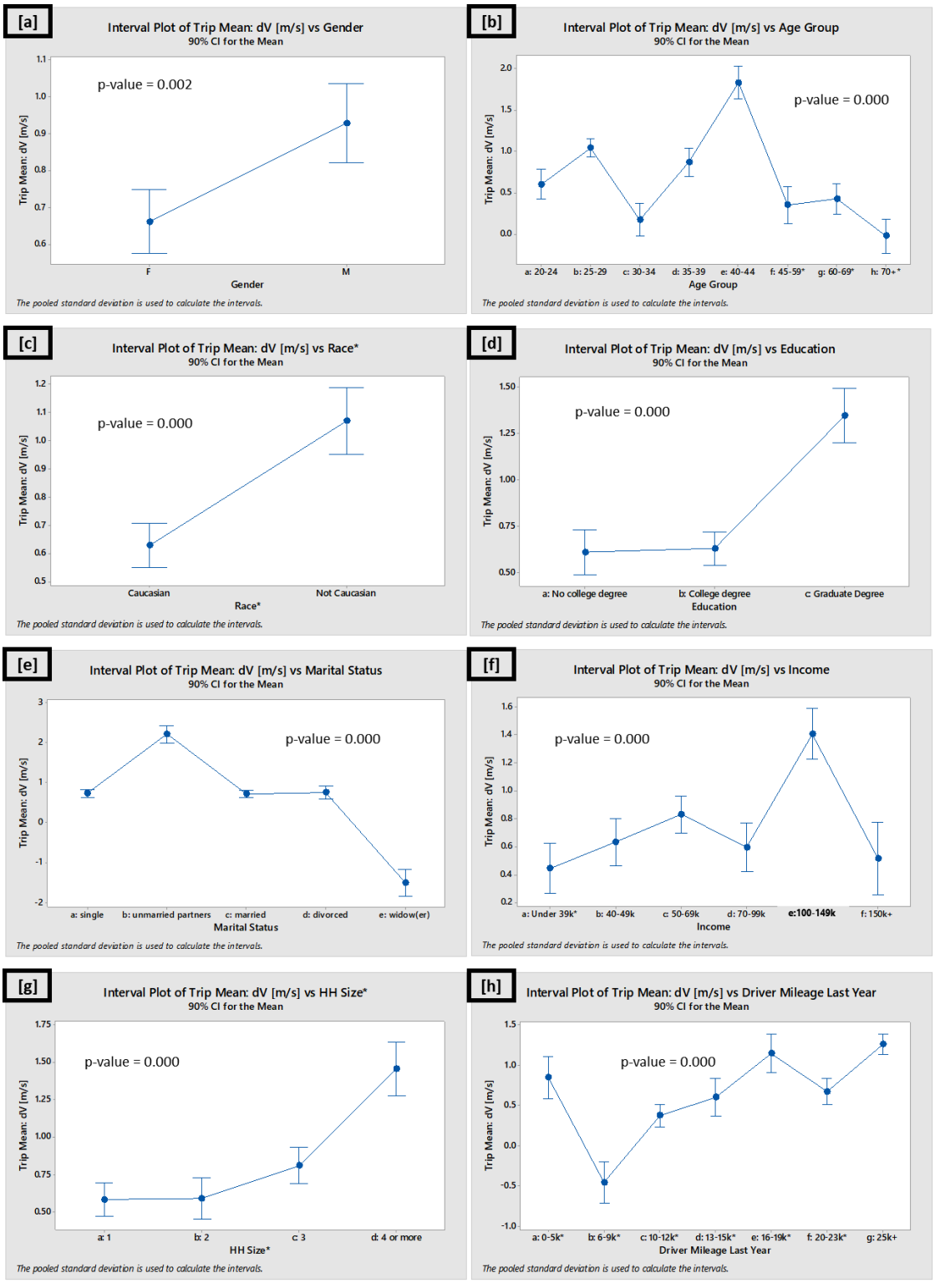


Figure 5.4 Mean Relative Velocity Segmented by Driver Attributes

Figure 5.5 shows how the average following distance of all car-following events comprising a complete trip varies according to driver attributes. All driver attributes exhibited statistically significant differences amongst the different subcategories of data. According to Figure 5.5a, female drivers maintain smaller inter-vehicle spacings than male drivers, on average. As illustrated in Figure 5.5b, driver age and inter-vehicle spacing are positively correlated, with younger drivers maintaining smaller gaps than older drivers. Figure 5.5e suggests that relative gap and marital status are strongly correlated.

According to Table A.5, the smallest average following distance was documented for the 25–29 age group (mean = 44.53 m). The largest average following distance was documented for the widow(er) marital status subcategory (mean = 83.77 m). The smallest within subcategory variation was calculated for the unmarried partners subcategory (standard deviation = 7.05 m), while the largest variation was recorded for the 6–9k driver mileage last year subcategory (standard deviation = 24.27 m). The smallest within attribute variation was calculated for the age attribute [mean(standard deviation) = 13.5 m]. The largest attribute variation occurred for the driver mileage last year attribute [mean(standard deviation) = 16.9 m].

Figure 5.6 documents how the average headway of all car-following events comprising a complete trip varies according to driver attributes. The only attributes that did not exhibit statistically significant differences in behavior across the subcategories are gender and race. As evident in Figure 5.6b, the youngest and oldest subcategories of drivers exhibit higher headways than middle-aged drivers; this is also corroborated by marital status in Figure 5.6e, where the widow(er) subcategory reports substantially higher headways than the rest of the sample.

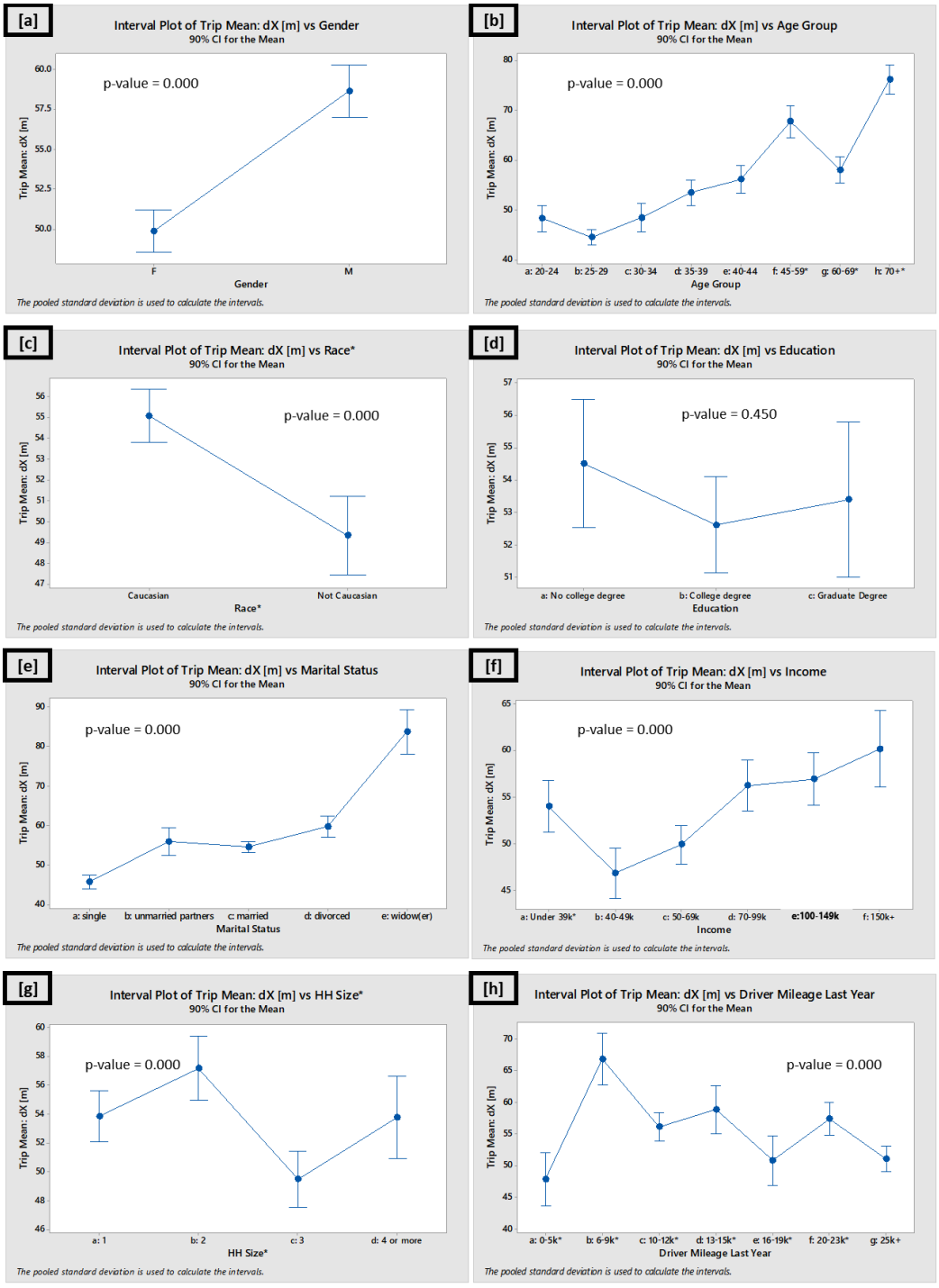


Figure 5.5 Mean Relative Following Distance Segmented by Driver Attributes

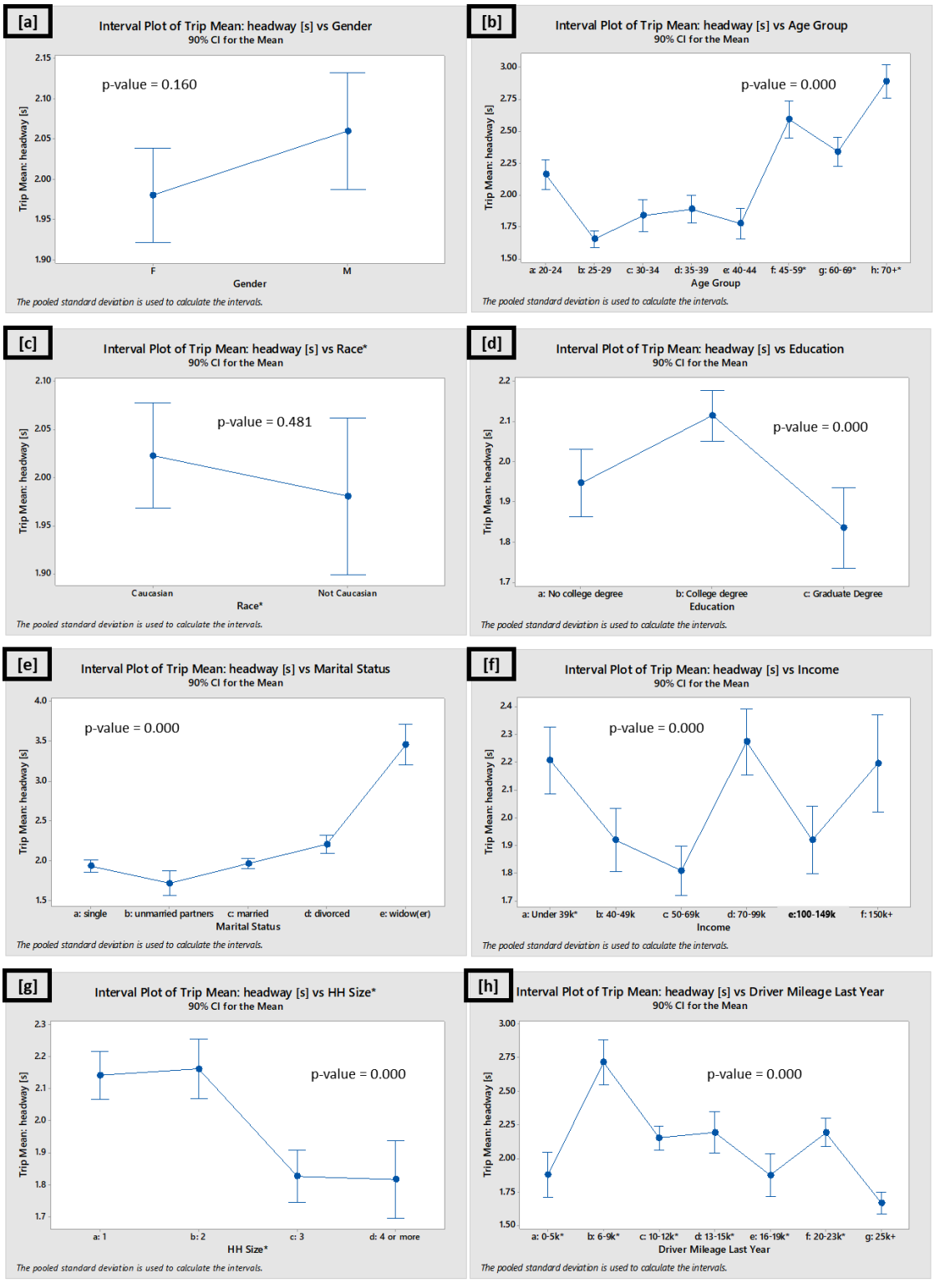


Figure 5.6 Mean Headway Segmented by Driver Attributes

According to Table A.6, the smallest headway was observed for the 25–29 age group (mean = 1.7 s), while the largest headway was observed for the widow(er) subcategory (mean = 3.4 s). The smallest within subcategory variation was observed for the unmarried partners subcategory of drivers (standard deviation = 0.2 s), while the largest variation was observed for the 6–9k driver mileage last year subgroup (standard deviation = 1.0s). The marital status and age attributes were calculated to have the smallest average standard deviation across all subcategories [mean(standard deviation) = 0.60 s], while the income attribute had the most dispersion across the subcategories comprising the attribute [mean(standard deviation) = 0.71 s].

5.2.2.2. Individual Parameter Results

Figure 5.7 shows how the W99 standstill distance (cc0) coefficient estimates vary across the eight driver attributes. The gender and driver mileage last year attributes were not found to have statistically significant differences across the subcategories. There are not clear, succinct qualitative trends visible in Figure 5.7a through Figure 5.7g.

As documented in Table A.7, the 30–34 age group reported the highest calibrated standstill distance coefficient estimate (mean = 5.7 m), while the $50–69k income subcategory reported the lowest calibrated standstill distance coefficient estimate (mean = 3.8 m). These results are a bit surprising, as one would have expected the widow(er) and the 25–29 age group to have reported the highest and lowest estimated value for standstill distance, as suggested by the trip summary statistics. The within subcategory standard deviations were remarkably similar across the attributes. The 13–15k subcategory for driver mileage last year reported the highest within subcategory variation (standard deviation = 2.8 m), while the 0–5k subcategory reported the lowest variation (standard

deviation = 2.2 m). The smallest average within attribute variation was calculated for the race attribute [mean(standard deviation = 2.38 m], while the driver mileage last year attribute reported the highest [mean(standard deviation = 2.48 m].

Figure 5.8 illustrates how the W99 spacing time (cc1) coefficient estimates vary across the subcategories of driver attributes. The gender, race, and education attainment attributes were not found to vary across the subcategories at a statistically significant level. There are not clear, succinct qualitative trends visible in Figure 5.8a through Figure 5.8g. The age, marital status, and miles driven last year attributes each had one subcategory with substantially higher spacing time estimated parameter coefficients compared to adjacent categories: the 70+ age group, widow(er), and 6–9k miles driven last year subcategories, respectively. It is worth noting that these are all categories that showed correlation in other Kruskal-Wallis ANOVA analyses (e.g., headway vs. driver attributes in Figure 5.6).

Quantitative results are documented in Table A.8. The widow(er) subcategory reported the highest calibrated spacing time value (mean = 2.16 s), while the 25–29 age group reported the lowest calibrated spacing time value (mean = 0.58 s). The widow(er) subcategory reported the highest within subcategory variation (standard deviation = 1.11 s). Reassuringly, these results are consistent with the average time headway summary statistics. The 25–29 age group reported the lowest within subcategory variation (standard deviation = 0.27 s). The age attribute achieved the lowest average within attribute variation [mean(standard deviation) = 0.49s] , while the marital status attribute reported the highest average variation [mean(standard deviation = 0.58 s].

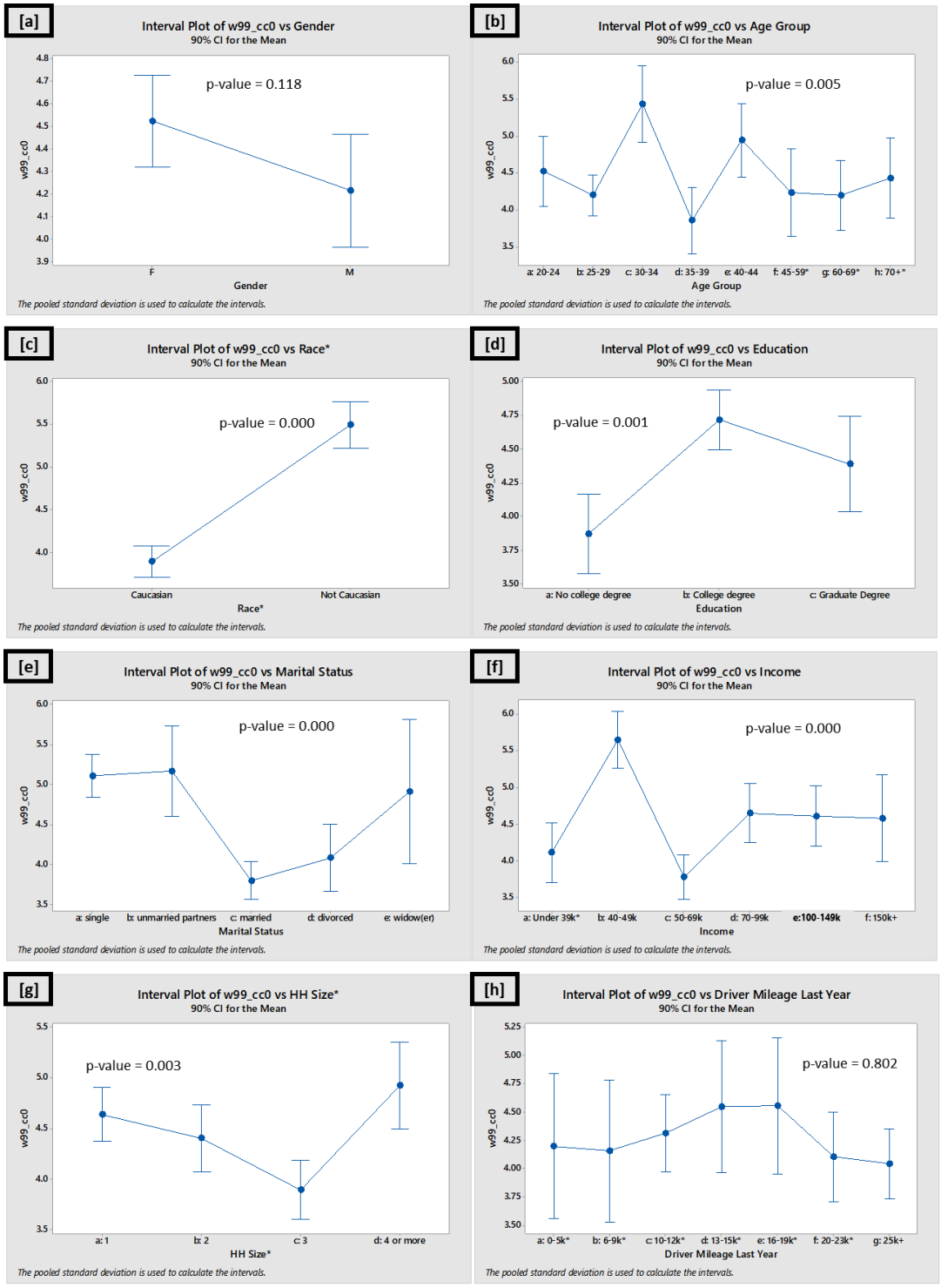


Figure 5.7 Estimated Wiedemann 99 Standstill Distance (cc0) Coefficient Segmented by Driver Attributes

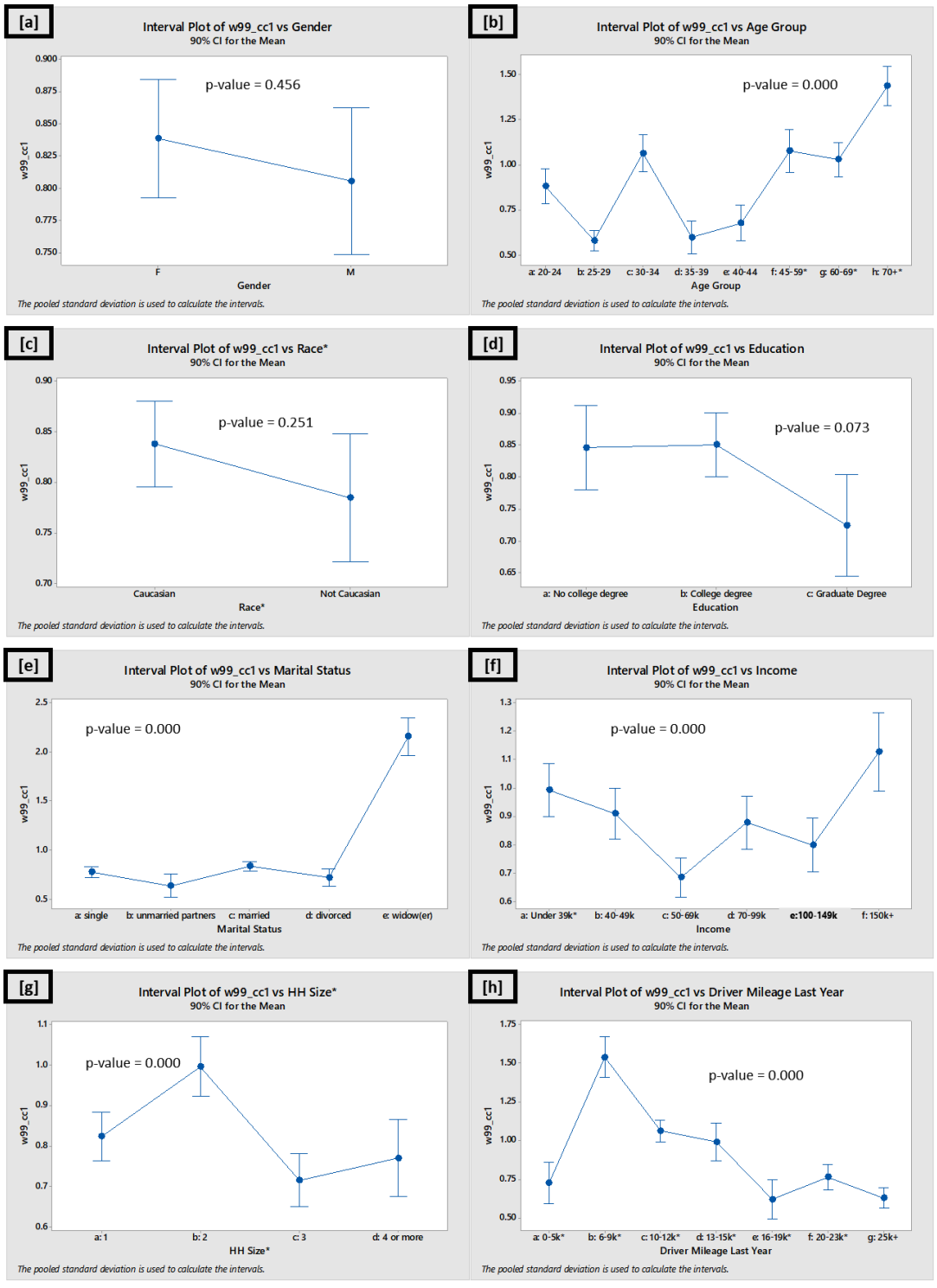


Figure 5.8 Estimated Wiedemann 99 Spacing Time (cc1) Coefficient Segmented by Driver Attributes

Figure 5.9 illustrates how the W99 following variation, or maximum drift, (cc2) coefficient estimates vary with driver attributes. The cc2 parameter defines the height of the following regime, which is defined by the SDXC and SDXO thresholds of the W99 model in Figure 3.2. This parameter represents the range of distances that a driver subconsciously aims to stay within while following a leading vehicle (the actual range is the cc2 estimated parameter ± the average following distance). Smaller estimated cc2 parameter coefficients suggests a smaller range of values that a driver 'drifts' between while following a leading vehicle (i.e., a more attentive or aggressive driver). The gender, educational attainment, household size, and driver mileage last year attributes did not vary across the subcategories at a statistically significant level for the following variation parameter.

As documented in Table A.9, the $40–49k income subcategory reported the highest average calibrated following variation coefficient (mean = 11.86 m), while the 40–59 age group subcategory reported the lowest average following variation estimated coefficient (mean = 8.13 m). The 35–39 age group reported the highest within subcategory variation (standard deviation = 5.53 m) for the following variation estimated coefficient, while the highest income category reported the lowest within subcategory variation (standard deviation = 3.75 m). The smallest within attribute variation was calculated for the income attribute [mean(standard deviation = 4.6 m), while the largest attribute variation occurred in the gender attribute [mean(standard deviation = 4.9 m).

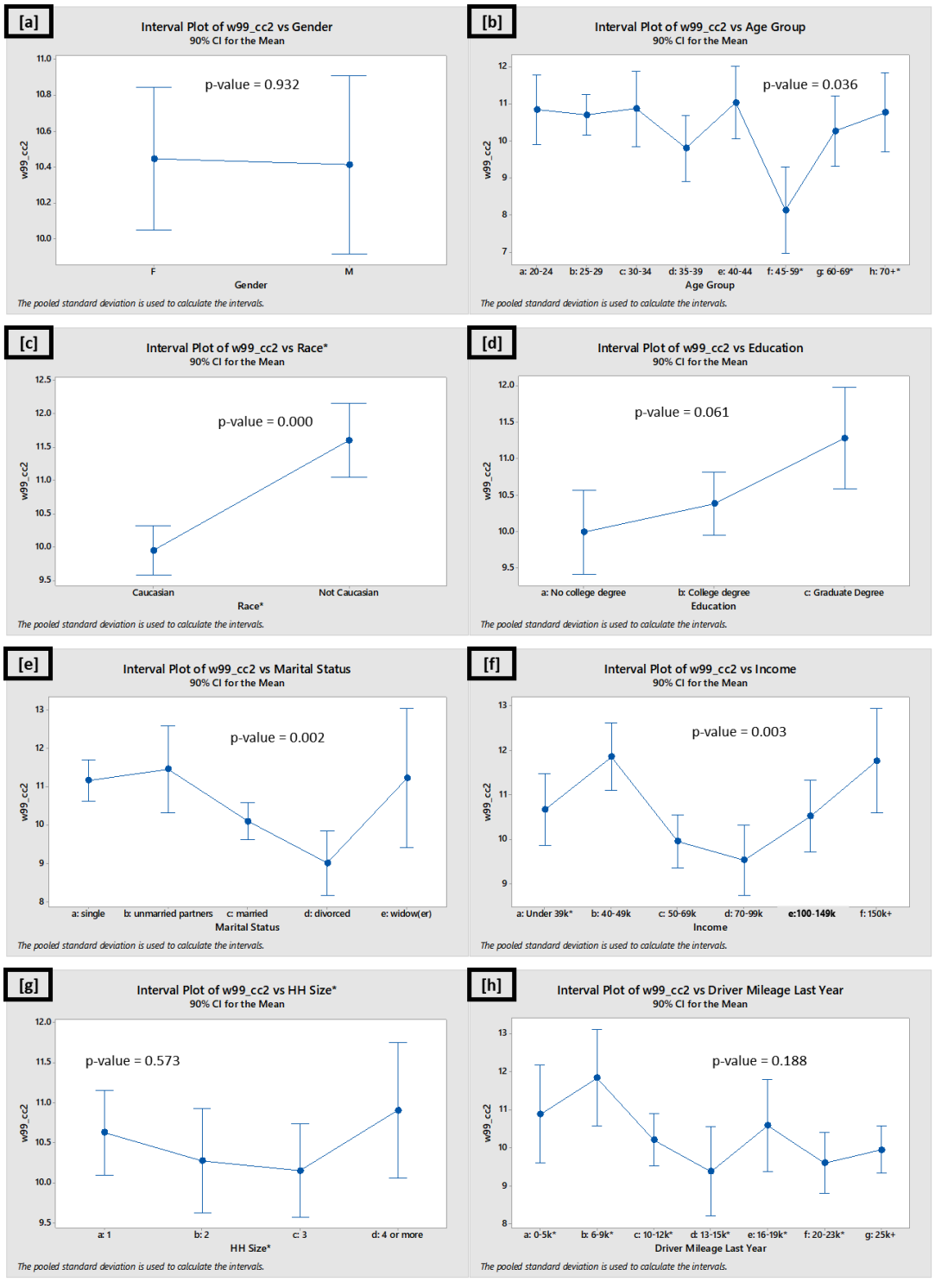


Figure 5.9 Estimated Wiedemann 99 Following Variance, Maximum Drift (cc2) Coefficient Segmented by Driver Attributes

Figure 5.10 depicts how the threshold for entering 'following' (cc3) coefficient estimates for the W99 model change across subcategories of driver attributes; this parameter is used in the calculation of the SDXV perception threshold in Figure 3.2, which separates the free flow and the approaching regimes of the W99 car-following model (i.e., as the magnitude of this parameter decreases, the perception-reaction threshold for approaching a lead vehicle increases and a driver reacts more quickly to the presence of a leading vehicle). The educational attainment, income, household size, and driver mileage last year attributes were not found to vary at a statistically significant level across the various subcategories.

As documented in Table A.10, the 13–15k miles driven last year subcategory reported the largest estimated value for the threshold for entering 'following' coefficient (mean = -23.6 s). The 45–59 age group reported the lowest threshold for entering 'following' calibrated coefficient (mean = -20.0 s). Moreover, the 13–15k miles driven last year subcategory reported the smallest within subcategory variation (standard deviation = 3.93 s), while the 45–59 age group reported the largest standard deviation within the subcategory (standard deviation = 6.96 s). The smallest average attribute variation was recorded for driver race [mean(standard deviation = 5.48 s), while the largest attribute variation was calculated for driver marital status [mean(standard deviation = 5.82 s].

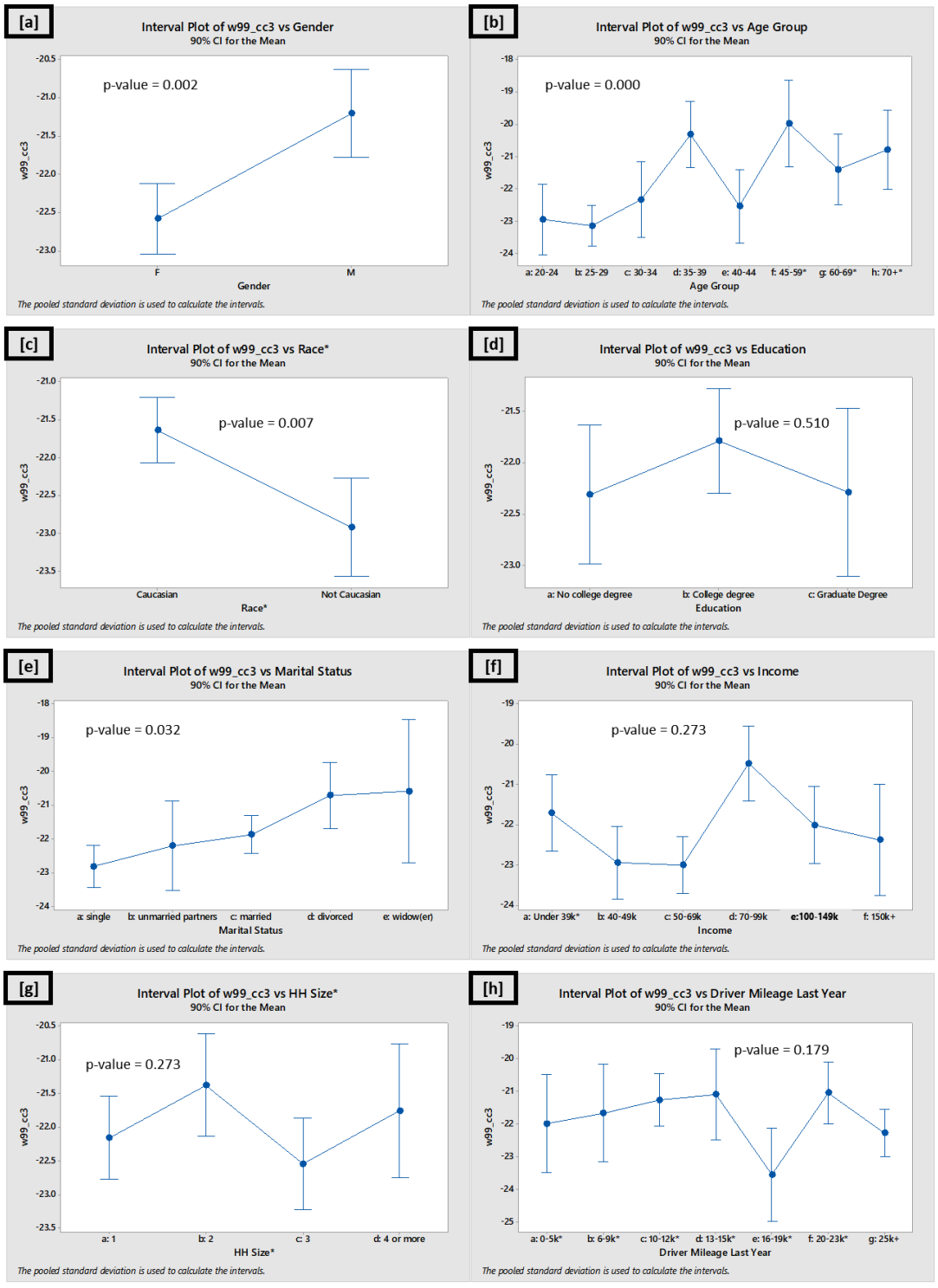


Figure 5.10 Estimated Wiedemann 99 Threshold for Entering Following (cc3) Coefficient Segmented by Driver Attributes

Figure 5.11 describes how the W99 negative following threshold (cc4) coefficient estimates vary according to driver attributes. The negative following threshold coefficient is used to calculate the SDVC and SDVO thresholds of the W99 model; these thresholds describe the range of relative velocity values a driver subconsciously tries to stay within while following a lead vehicle. Specifically, the negative following threshold dictates how quickly the following vehicle reacts when approaching the lead vehicle (i.e., the larger the absolute value of the following threshold, the greater the acceptable difference in relative velocity). The educational attainment, income, and household size attributes were not found to vary at a statistically significant level across the subcategories. Trends across the subcategories are observed for the gender (Figure 5.11a), age group (Figure 5.11b), race (Figure 5.11c), and marital status (Figure 5.11e) attributes. Figure 5.11a suggests that the average male negative following threshold is larger than that of female drivers. Figure 5.11b suggests that the negative following threshold is correlated with driver age; that is, as age increases, the negative following threshold generally increases. Moreover, Figure 5.11e indicates that marital status and negative following threshold estimated coefficients are correlated.

According to Table A.11, the widow(er) subcategory reported the most extreme average calibrated negative following threshold value (i.e., the least reactive to changes in relative velocity) (mean = -1.7 m/s) and the highest within subcategory variation (standard deviation = 1.8 m/s). The unmarried partners subcategory reported the most modest calibrated value for the negative following threshold coefficient (i.e., more sensitivity to changes in relative velocity) (mean = -0.25 m/s) and the lowest within subcategory variation (standard deviation = 0.52 m/s). The race attribute reported the lowest average variation across all subcategories comprising the attribute [mean(standard deviation) = 1.1 m/s], while the driver mileage last year attribute reported the highest variation

[mean(standard deviation) = 1.3 m/s].

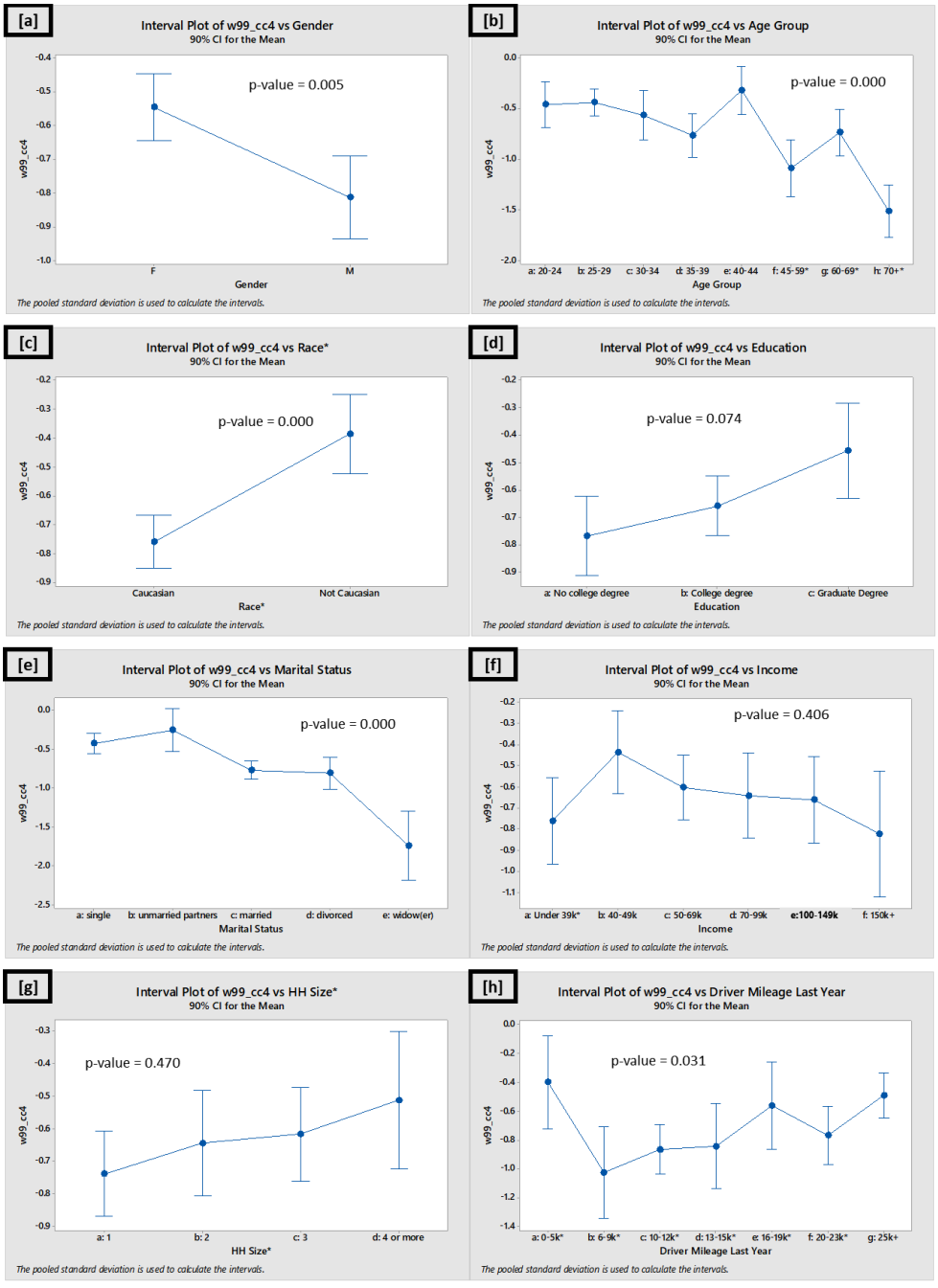


Figure 5.11 Estimated Wiedemann 99 Negative Following Threshold (cc4) Coefficient Segmented by Driver Attributes

Figure 5.12 portrays how the W99 positive following threshold (cc5) estimated coefficients vary across driver attributes. The positive following threshold coefficient is used to calculate the SDVC and SDVO thresholds of the W99 model, as shown in Figure 3.2; these thresholds describe the range of relative velocity values a driver subconsciously tries to stay within while following a leading vehicle. The positive following threshold dictates how quickly the following vehicle reacts when separating from the lead vehicle (i.e., the larger the absolute value of the following threshold, the greater the acceptable difference in relative velocity). The gender, educational attainment, and income attributes were not found the vary across subcategories at a statistically significant level. Figure 5.12e suggests that marital status and the calibrated positive following threshold are correlated.

Quantitative results are documented in Table A.12. The widow(er) and unmarried partners subcategories reported the highest (mean = 2.05 m/s) and lowest (mean = 0.95 m/s) calibrated coefficient estimates, respectively; this is consistent with the trends observed with the negative following threshold estimated coefficients. The 25–29 age group reported the lowest within subcategory variation (standard deviation = 1.02 m/s), while the 45–59 age group reported the highest (standard deviation = 1.61 m/s). The education attainment attribute reported the lowest average variation within subcategories comprising an attribute [mean(standard deviation) = 1.25 m/s], while the age and income attributes reported the highest average variation [mean(standard deviation) = 1.32 m/s].

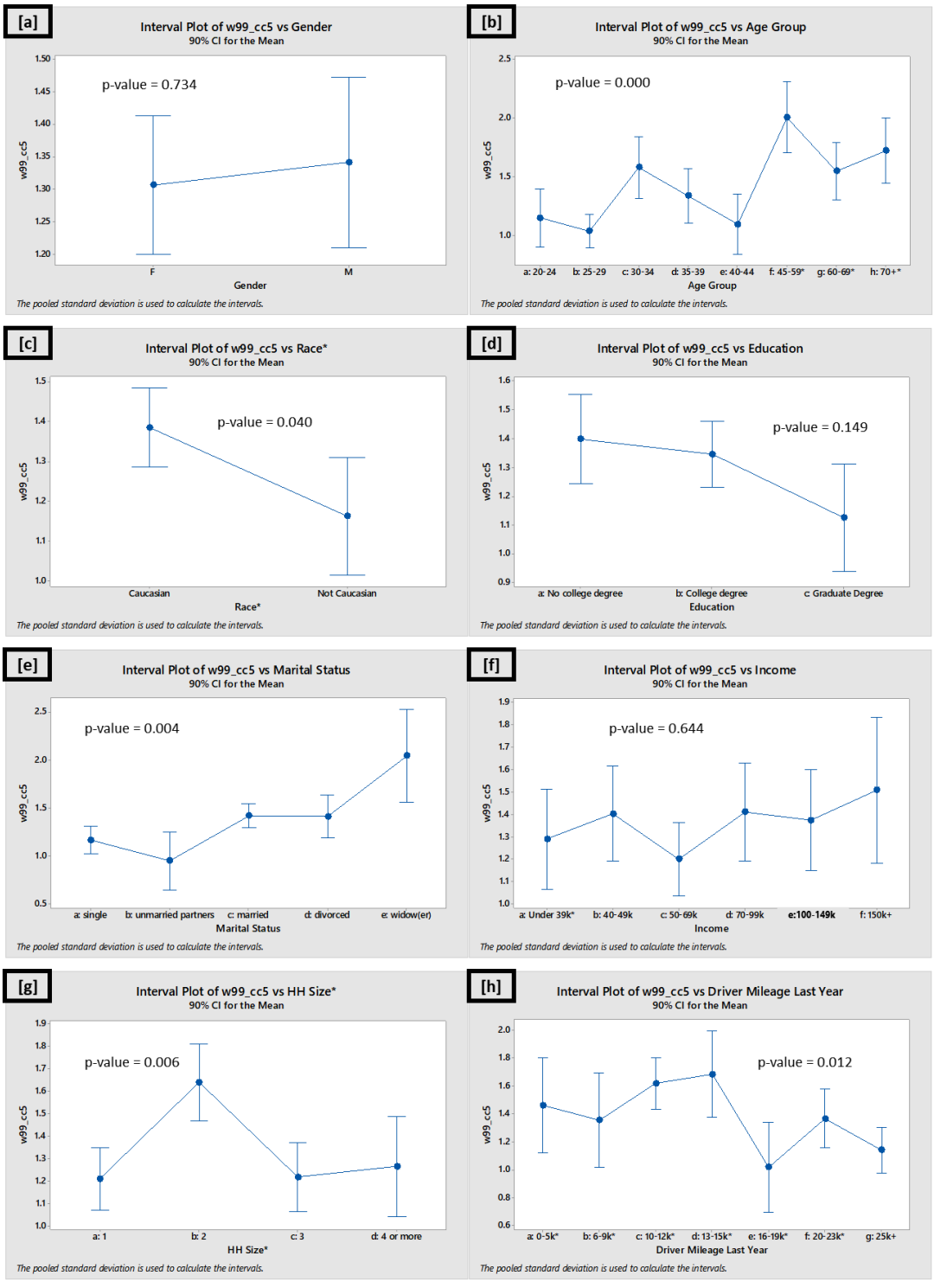


Figure 5.12 Estimated Wiedemann 99 Positive Following Threshold (cc5) Coefficient Segmented by Driver Attributes

Figure 5.13 illustrates how the W99 speed dependency of oscillation (cc6) estimated coefficients vary with driver attributes; this parameter is also used in the calculation of the SDVC and SDVO thresholds in Figure 3.2. The speed dependency of oscillation parameter specifically impacts the magnitude of the oscillation of driving behavior on a psychophysical plane (i.e., a larger magnitude represents wider oscillation, while a smaller magnitude represents narrower oscillation). Education attainment, income, household size, and reported driver mileage last year were all found to not vary at a statistically significant level across the subcategories comprising the driver attributes. Figure 5.13a indicates that male drivers have a higher speed dependency of oscillation parameter coefficient than female drivers. Moreover, it appears that the speed dependency of oscillation estimated calibration coefficient is positively correlated with driver age, with the exception of a large spike in this value for the 35–39 age group.

According to Table A.13, the divorced subcategory was estimated to have the highest average speed dependency of oscillation coefficient (i.e., wider oscillation behavior) (mean = 3.3 x 10^{-4} rad/s); this subcategory also had the highest within subcategory variation. The 20–24 age group was estimated to have the smallest speed dependency of oscillation coefficient (i.e., narrower oscillatory behavior) (mean = 1.7 x 10^{-4} rad/s); the 40–44 age group achieved the smallest within subcategory variation. The income attribute reported the smallest average variation across all subcategories comprising the attribute (2.2 x 10^{-4} rad/s). The driver mileage last year attribute reported the largest average variation across all subcategories (2.3 x 10^{-4} rad/s).

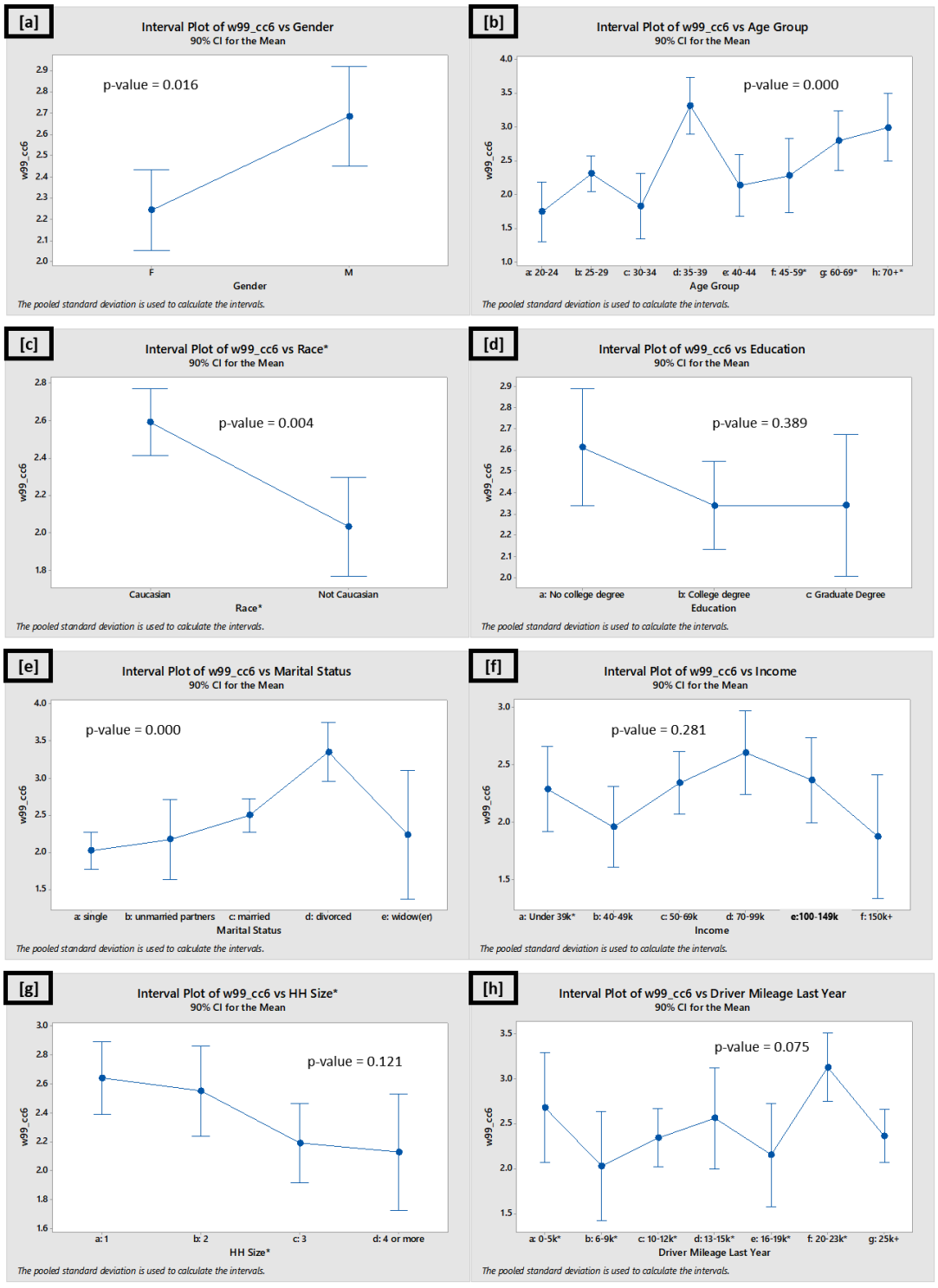


Figure 5.13 Estimated Wiedemann 99 Speed Dependency of Oscillation (cc6) Coefficient Segmented by Driver Attributes

Figure 5.14 describes how the W99 oscillation acceleration (cc7) parameter estimates vary with driver attributes; this parameter represents the maximum acceleration of the following vehicle during the oscillatory behavior of the following process. Only three of the eight driver attributes were found to vary across subcategories at a statistically significant level: age group, race, and income. The 40–44 age group was estimated to have the lowest oscillatory acceleration (mean = 0.62 m/s^2), as well as the lowest within subcategory variation (standard deviation = 0.8 m/s^2), as documented in Table A.14. The $40–49k income subcategory was estimated to have the highest oscillatory acceleration (mean = 2.13 m/s^2). The widow(er) subcategory had the highest within subcategory variation (standard deviation = 3.0 m/s^2). The income attribute achieved the lowest average variation across all subcategories (standard deviation = 1.8 m/s^2). Lastly, the marital status attribute had the highest average variation across all subcategories (standard deviation = 2.1 m/s^2); it is worth noting that the widow(er) subcategory significantly increased the variation across the marital status attribute.

Figure 5.15 describes how the W99 standstill acceleration (cc8) coefficient estimates vary as a function of subcategories of driver attributes; this parameter represents the maximum acceleration of the following vehicle when accelerating from a stop. None of the attributes were found to vary at a statistically significant level across the different subcategories. Again, this could be attributable to data incompleteness, where not all trips consist of data collected at very low speeds.

Figure 5.16 describes how the W99 acceleration at 80 kph (cc9) coefficient estimates vary with driver attributes; this parameter represents the maximum acceleration of the following vehicle when it is travelling at high speeds. Race was the only attribute for which there were statistically significant differences in the calibrated cc9 values across the subcategories.

Practically speaking, the negative inverse of the cc7, cc8, and cc9 parameters are used to calculate the maximum deceleration rates of the W99 model. It is quite surprising that so little valuable differences between the subcategories of driver attributes are observed, especially given the variation of the observed minimum, maximum, and average acceleration values in the original SHRP2 data, as shown in Figure 5.1 through Figure 5.3. The author has two hypotheses for this discrepancy. The first hypothesis is concerned that the W99 model may oversimplify the nuanced differences in acceleration behavior evident in the naturalistic data. However, this observation is more likely an artifact of the calibration process itself. The W99 car-following model is the most complex to calibrate, given the necessity to calibrate eleven separate parameters. It has been observed through other calibration efforts, such as Methods and tolls for supporting the Use calibration and validation of Traffic simulation moDEls (MULTITUDE) project (European Cooperation in Science and Technology, n.d.), that the car-following model calibration problem likely does not produce a unique solution. The author is confident in the calibration process, as the genetic algorithm consistently converges with an acceptable root mean squared error between the predicted and observed spacing profiles. It is possible that the algorithm is leveraging the large number of calibration coefficients to its advantage (i.e., the algorithm identifies reasonable values for coefficients that are harder to calibrate and toys with the values for the remaining parameters to achieve a more optimal solution). In future research, the author recommends innovative calibration strategies, such as those documented in Punzo, Montanino, and Ciuffo (2015), to reduce the calibration complexity by only calibrating the most critical car-following model parameters.

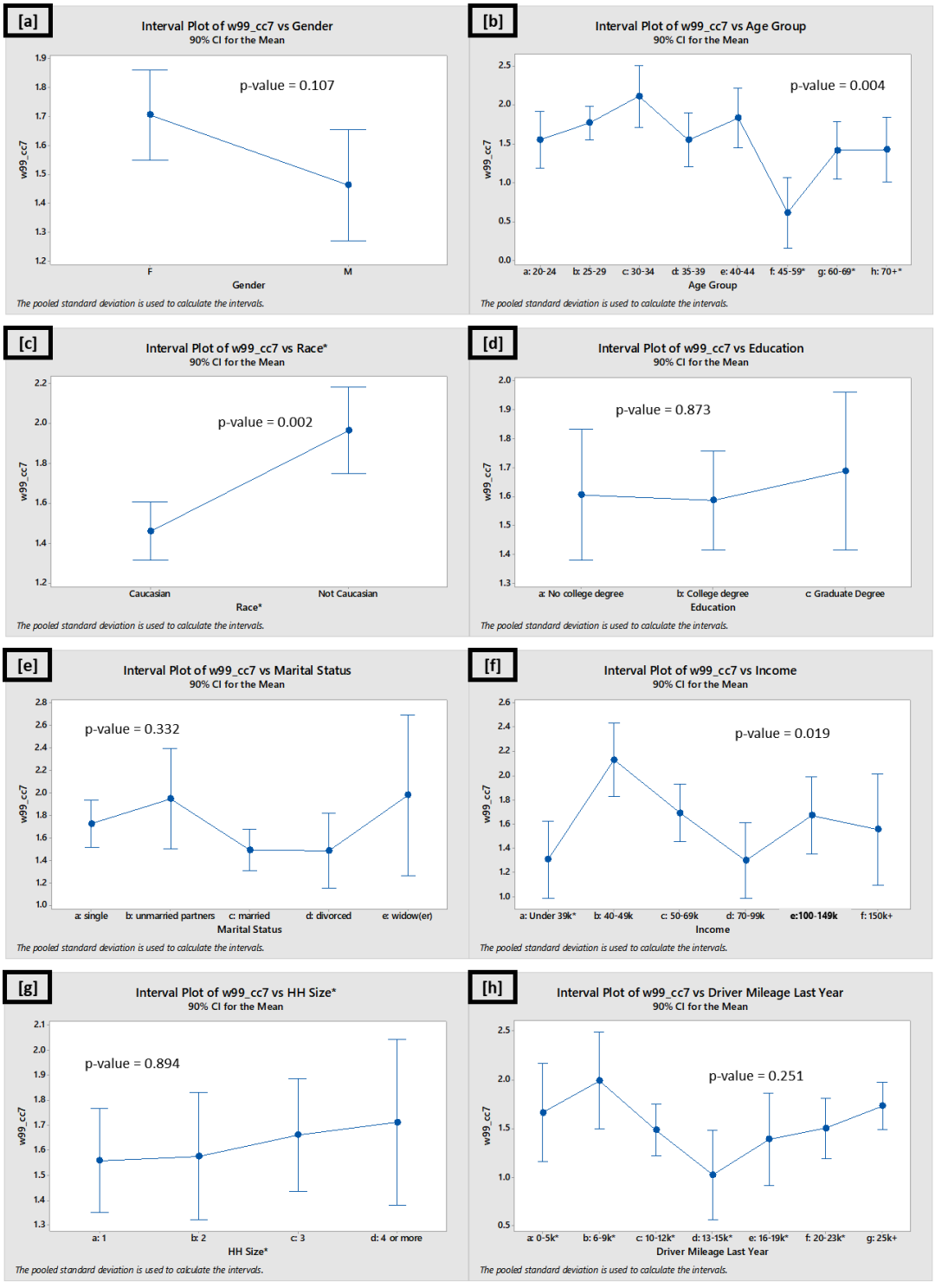


Figure 5.14 Estimated Wiedemann 99 Oscillation Acceleration (cc7) Coefficient Segmented by Driver Attributes

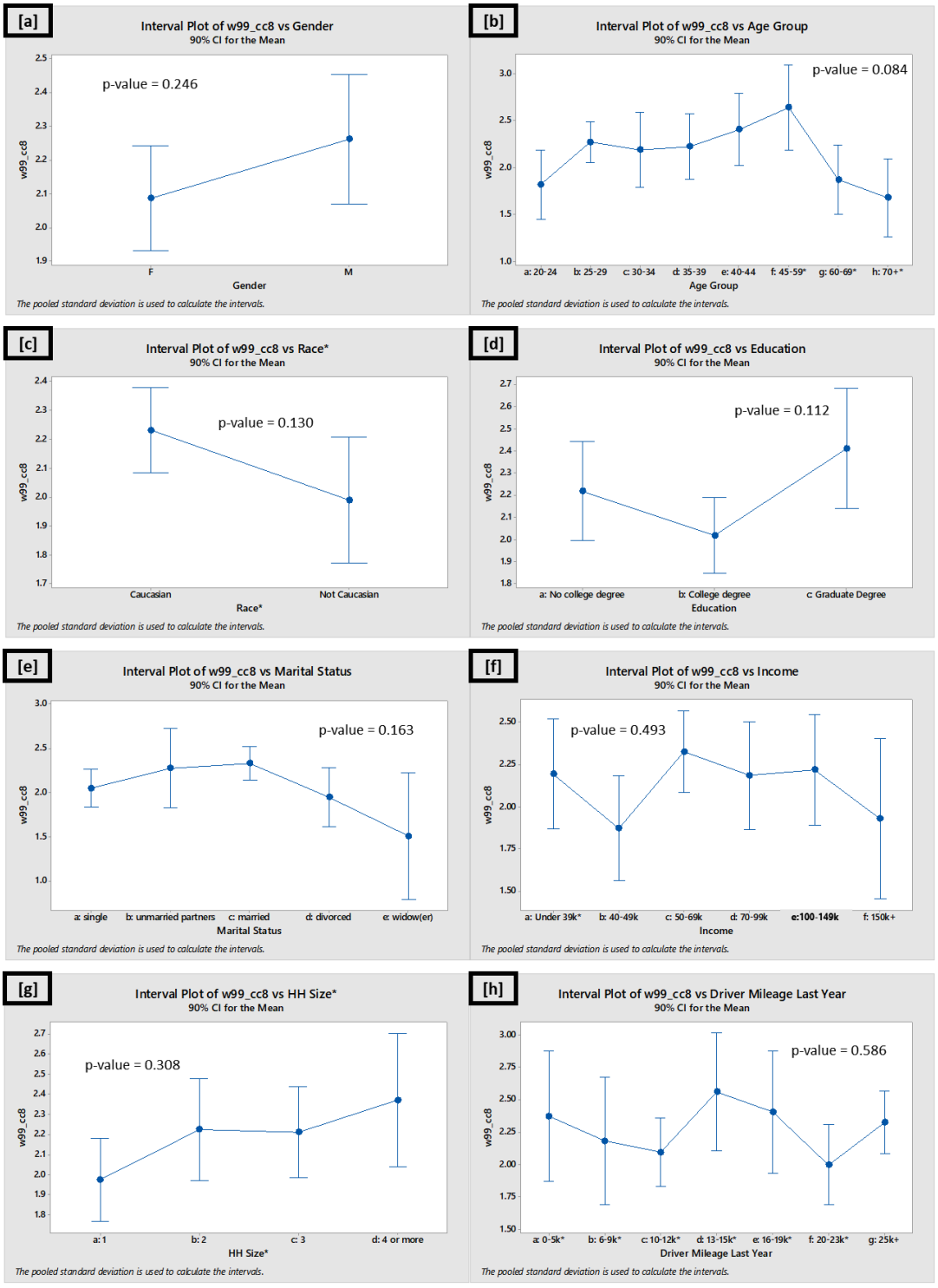


Figure 5.15 Estimated Wiedemann 99 Standstill Acceleration (cc8) Coefficient Segmented by Driver Attributes

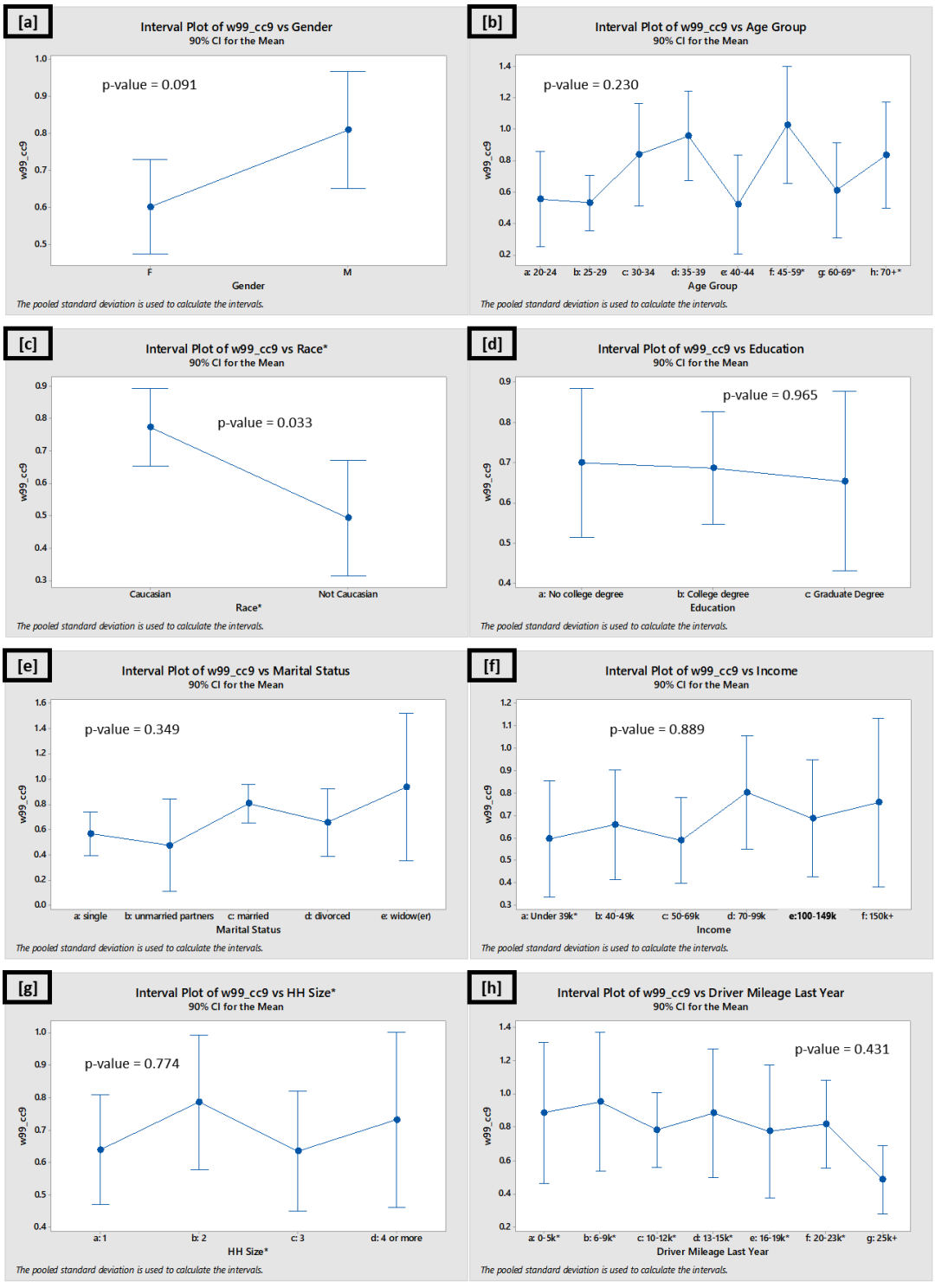


Figure 5.16 Estimated Wiedemann 99 Acceleration at 80 kph (cc8) Coefficient Segmented by Driver Attributes

Figure 5.17 describes how the calibrated desired velocity parameter of the W99 car-following model varies according to driver attributes. All driver attributes were found to vary at a statistically significant level across the different subcategories. As illustrated in Figure 5.17a, male drivers were estimated to have a lower desired velocity than female drivers. Moreover, in Figure 5.17b it is shown that age and desired velocity appear to have a negative correlation, with the exception of a large spike in desired velocity for the 40–44 age group.

According to Table A.17, the unmarried partners subcategory of the marital status attribute was estimated to have the highest desired velocity value (mean = 35.3 m/s), as well as the smallest within subcategory variation (standard deviation = 1.5 m/s). The widow(er) subcategory was estimated to have the smallest desired velocity value (mean = 27.3 m/s); this subcategory also had the highest within subcategory variation (standard deviation = 3.9 m/s). The age attribute had the lowest average variation across all of the subcategories comprising the attribute [mean(standard deviation) = 2.54 m/s], while the education attainment attribute appeared to have the highest variation across the subcategories [mean(standard deviation) = 2.96 m/s].

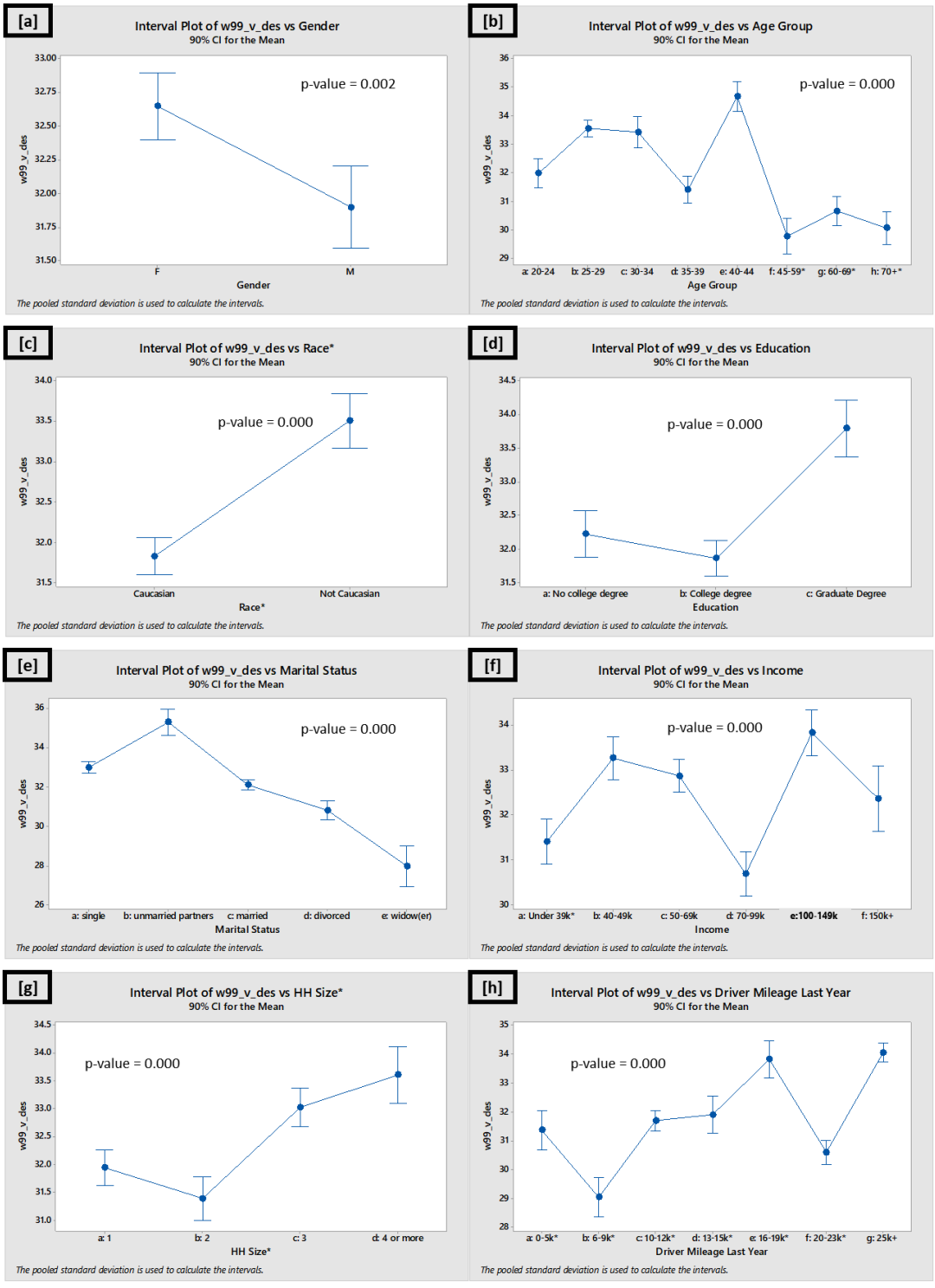


Figure 5.17 Estimated Wiedemann 99 Desired Velocity (v_des) Coefficient Segmented by Driver Attributes

Figure 5.18 illustrates how the Gipps desired acceleration parameter estimate varies across the eight driver attributes available with the NDS dataset. All driver attributes were found to vary at a statistically significant level across subcategories, except the driver mileage last year attribute. As illustrated in Figure 5.18a, the calibrated estimate for desired acceleration is smaller for male drivers than female drivers; moreover, Figure 5.18c indicates that the calibrated value for desired acceleration of Caucasian drivers is smaller than that of drivers that did not identify as Caucasian. Figure 5.18b indicates that the calibrated desired acceleration estimate is inversely correlated with driver age: as driver age increases, the desired acceleration coefficient estimate decreases. Lastly, Figure 5.18e suggests that driver desired acceleration is associated with reported marital status.

As documented in Table A.18, the largest average desired acceleration value was estimated for the $40–49k income subcategory (mean = 1.8 m/s^2). The smallest average desired acceleration value was estimated for the widow(er) marital status subcategory (mean = 0.6 m/s^2); this subcategory also achieved the smallest within subcategory variation (standard deviation = 0.55 m/s^2). The largest within subcategory variation was observed in the 30–34 age group (standard deviation = 1.18 m/s^2). The average attribute variation was calculated by averaging the standard deviation across all of the subcategories that comprise an attribute (e.g., the male and female subcategories comprise the gender driver attribute); the marital status attribute achieved the lowest attribute variation [mean(standard deviation) = 0.82 m/s^2]. The driver mileage last year attribute had the largest attribute variation [mean(standard deviation) = 0.962 m/s^2].

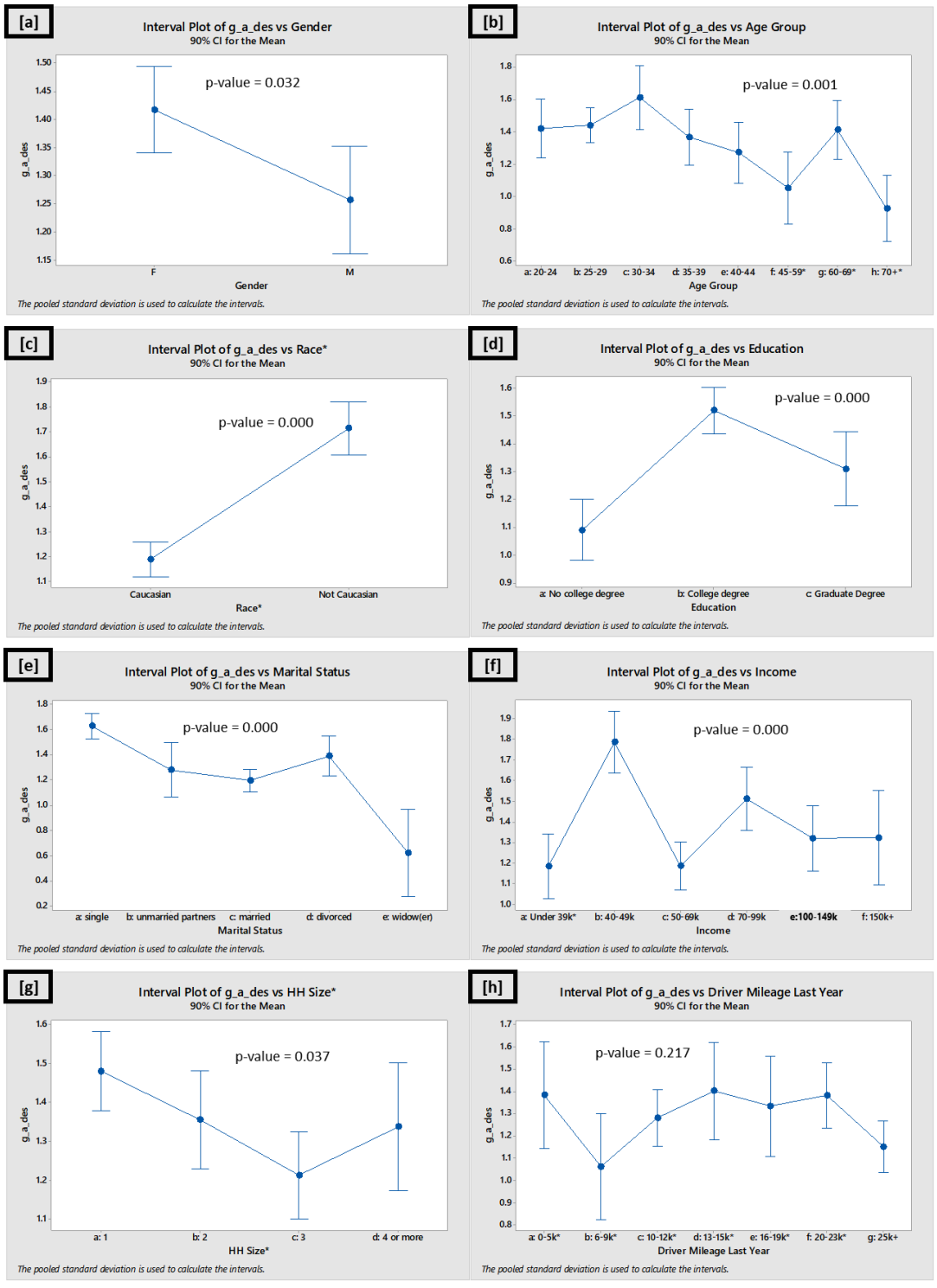


Figure 5.18 Estimated Gipps Desired Acceleration (a_des) Coefficient Segmented by Driver Attributes

Figure 5.19 depicts how the Gipps desired deceleration coefficient estimate varies according to driver attributes. The estimated parameter coefficients were found to vary at a statistically significant level for all available driver attributes. Figure 5.19a indicates that male drivers were estimated to have larger desired decelerations than female drivers. Figure 5.19c indicates that Caucasian drivers were estimated to have larger desired decelerations than drivers that reported that they are not Caucasian. Figure 5.19b suggests that there is a relationship between the estimated desired deceleration and driver age: middle age drivers were estimated to have larger desired deceleration values than younger and older drivers. Lastly, Figure 5.19g suggests that household size and desired deceleration rates are positively correlated: as household size grows, desired deceleration increases.

Quantitative results are provided in Table A.19. The smallest average desired deceleration value was estimated for the widow(er) subcategory (mean = -1.9 m/s^2); this subcategory also was observed to have the largest within subgroup variation (standard deviation = 1.21 m/s^2). The largest average desired acceleration value was estimated for the unmarried partners subcategory (mean = -3.5 m/s^2); this subcategory was also observed to have the smallest within subcategory variation (standard deviation = 0.54 m/s^2). The household size attribute had the smallest average attribute variation [mean(standard deviation) = 0.88 m/s^2], while the driver mileage last year attribute had the largest attribute variation [mean(standard deviation) = 0.94 m/s^2].

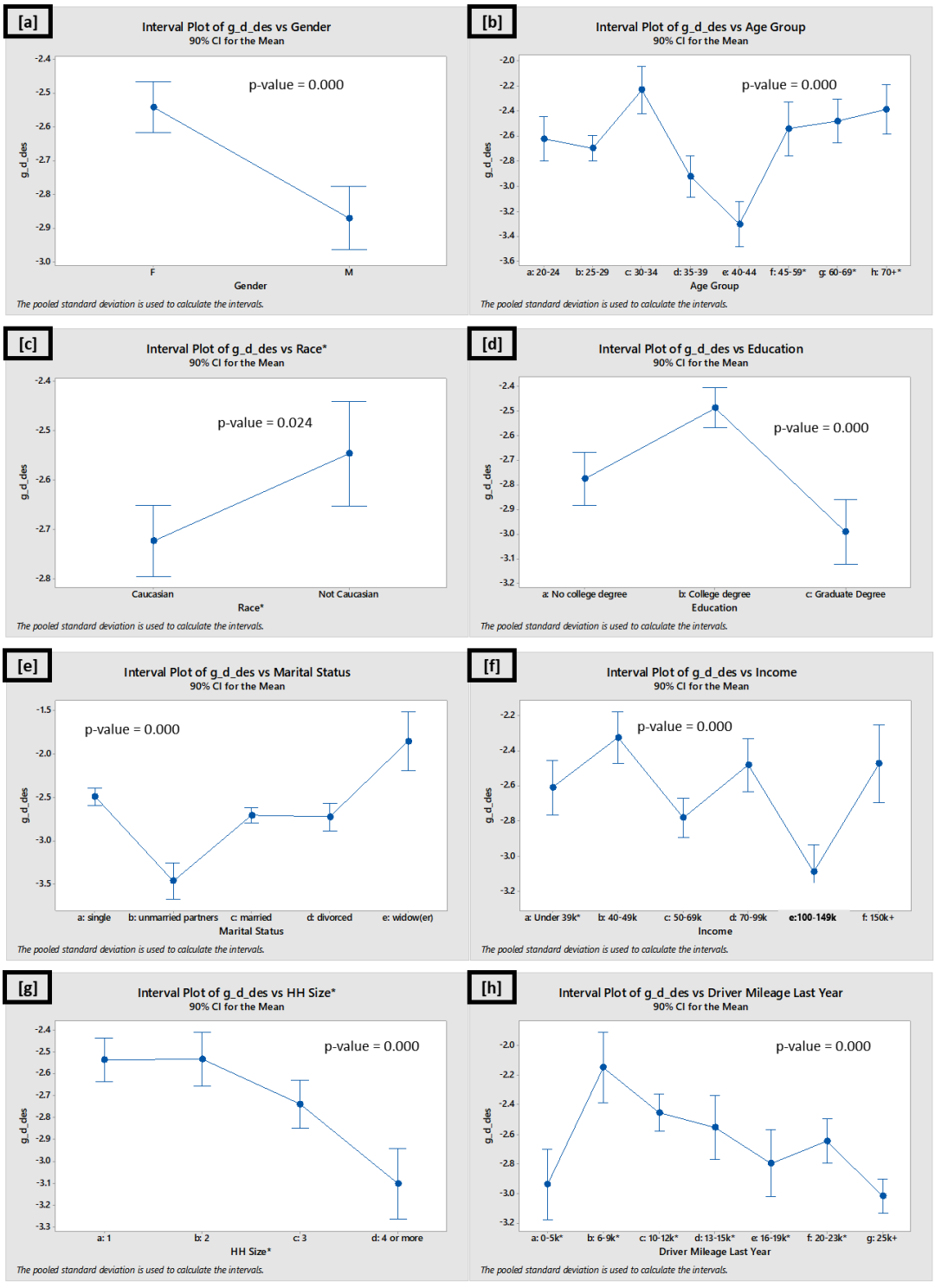


Figure 5.19 Estimated Gipps Desired Deceleration (d_des) Coefficient Segmented by Driver Attributes

Figure 5.20 shows how the Gipps parameter that captures the following vehicle driver's perception of the leading vehicle desired deceleration varies with driver specific attributes. The estimated parameter coefficients were found to vary across subcategories of driver attributes at a statistically significant level. Figure 5.20a indicates that male drivers were estimated to have smaller perceptions of leading vehicle desired decelerations than female drivers. Figure 5.20b suggests that there is a relationship between the estimated desired deceleration and driver age: middle age drivers were estimated to have smaller perceptions of leading vehicle desired deceleration values than younger and older drivers. If the trends in the data sound familiar, it is because we observed these trends with a previously described parameter: desired deceleration; in fact, the relationship between the estimated perceived driver deceleration rate and driver attributes very closely follows the relationships observed for the estimated following vehicle desired deceleration rate and driver attributes. This is because the following vehicle desired deceleration parameter and the following vehicle's perception of the leading vehicle's desired deceleration parameter are highly correlated (Pearson R correlation coefficient = 0.86) (Hammit, Ghasemzadeh, et al., 2018).

As documented in Table A.20, the smallest average perception of leader's desired deceleration value was estimated for the widow(er) subcategory (mean = -1.7 m/s^2); this subcategory was observed to have the largest within subgroup variation (standard deviation = 1.1 m/s^2). The largest average perception of leading vehicle desired acceleration coefficient was estimated for the unmarried partners subcategory (mean = -3.1 m/s^2); this subcategory was also observed to have the smallest within subcategory variation (standard deviation = 0.7 m/s^2). The household size attribute had the smallest average attribute variation [mean(standard deviation) = 0.90 m/s^2], while the driver mileage last year

attribute was found to have the largest attribute variation [mean(standard deviation) = 0.94 m/s^2].

Figure 5.21 illustrates how the Gipps desired distance at a stop (g_min) estimated parameter coefficient varies across subcategories of driver attributes. The desired distance at a stop parameter varied at a statistically significant level for half of the available driver attributes: age, race, marital status, and income.

According to Table A.21, the smallest average desired distance at a stop parameter was estimated for the 35–39 age group subcategory (mean = 3.2 m). The largest average desired distance at a stop parameter was estimated for the widow(er) subcategory of the marital status attribute (mean = 5.9 m). The widow(er) subcategory was one of three to be estimated to have desired distance at a stop parameter status above 5m; the other two subcategories were 60–69 and 70+ age groups. The smallest within subcategory variation was observed for the $40–49k income group (standard deviation = 2.6 m). The largest within subcategory variation was observed for the widow(er) subcategory (3.7 m). The smallest average variation within an attribute was observed for the race attribute [mean(standard deviation) = 2.98 m]. The largest average variation within an attribute was observed for the driver mileage last year [mean(standard deviation) = 3.6 m].

Figure 5.22 portrays how the estimated Gipps driver reaction time (t_rxn) coefficient estimate varies across driver attributes. The estimated reaction time coefficient varied at a statistically significant level for all attributes except driver gender.

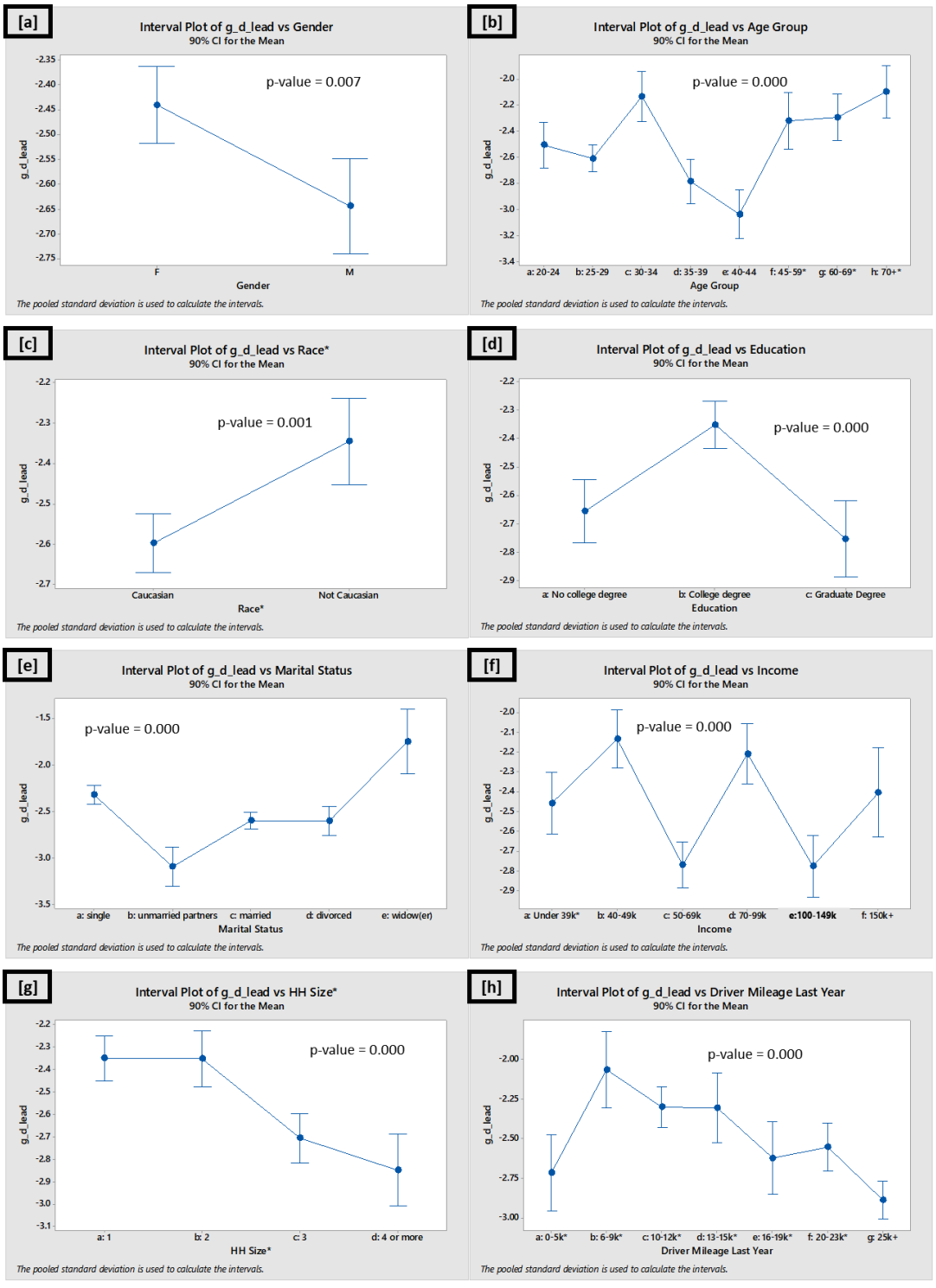


Figure 5.20 Estimated Gipps Perceived Desired Deceleration of Leading Vehicle (d_lead) Coefficient Segmented by Driver Attributes

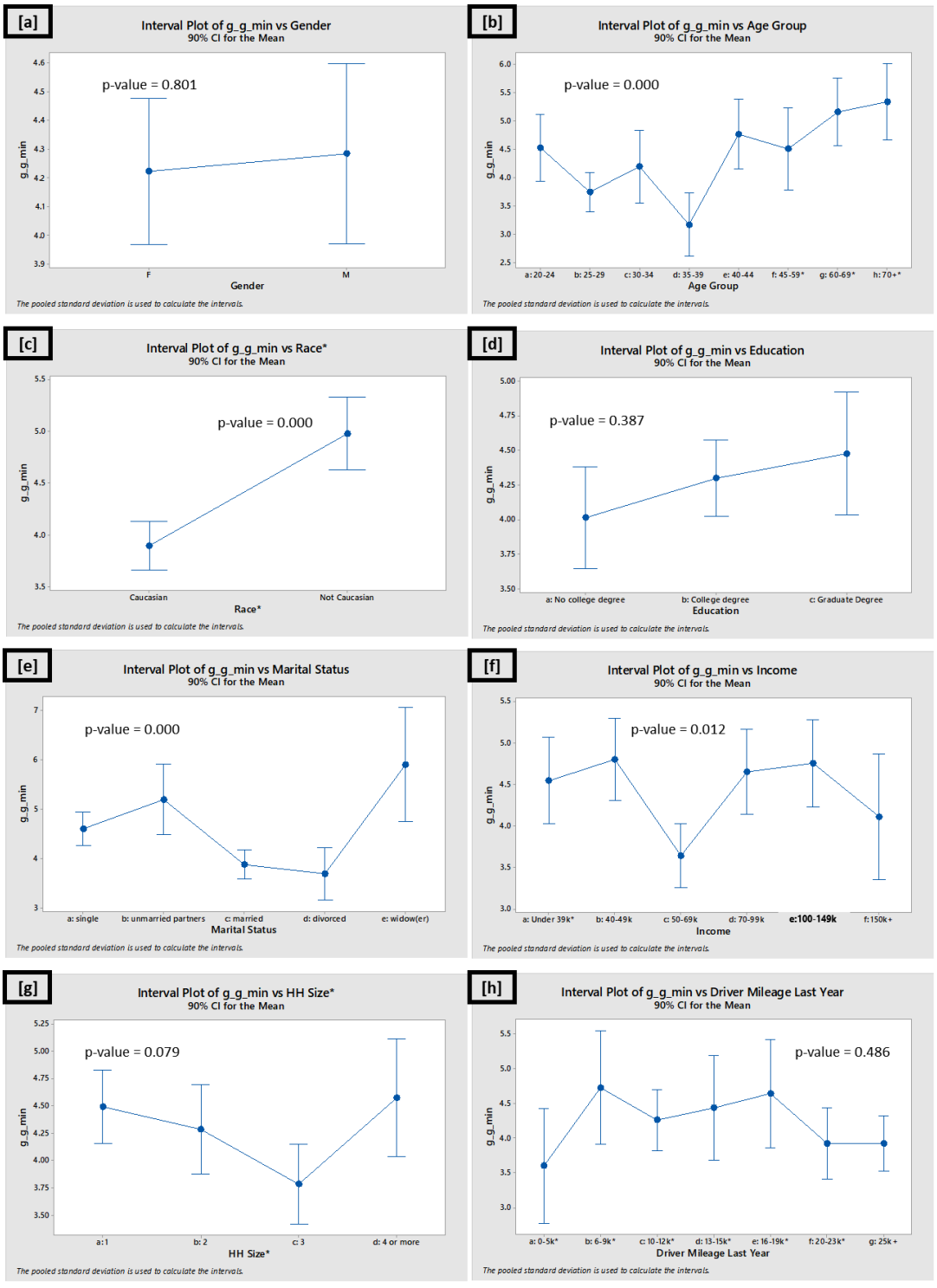


Figure 5.21 Estimated Gipps Desired Gap at a Stop (g_min) Coefficient Segmented by Driver Attributes

According to Figure 5.22 and Table A.22, the smallest average reaction time parameter coefficient was estimated for the $50–69k subcategory of the income attribute (mean = 0.50 s); this subcategory was also found to have the smallest within subcategory variation (standard deviation = 0.40 s). The largest average reaction time coefficient was estimated for the widow(er) subcategory of the marital status driver attribute (mean = 1.0 s); this subcategory was also found to have the largest within subcategory variation (standard deviation = 0.60 s). The educational attainment attribute had the smallest average variation across the subcategories [mean(standard deviation) = 0.50 s]. The age attribute was found had the largest average variation across the comprising subcategories [mean(standard deviation) = 0.52 s].

Figure 5.23 illustrates how the Gipps' desired velocity (v_des) estimated coefficient varies across driver specific attributes. The desired velocity estimates were found to vary across all subcategories of driver attributes at a statistically significant level. In Figure 5.23a, it is illustrated that the estimated value for desired velocity is smaller for male drivers than it is for female drivers. In Figure 5.23c, it is observed that the estimated parameter coefficient for desired velocity is smaller for Caucasian drivers than for drivers that do not identify as Caucasian. Figure 5.23b suggests that desired velocity and driver age are inversely correlated excluding the 40–44 age group, which is significantly higher than adjacent subcategories. Figure 5.23f indicates that driver income and desired velocity may be positively correlated apart from the $70–99k subcategory, which is abruptly smaller than adjacent subcategories.

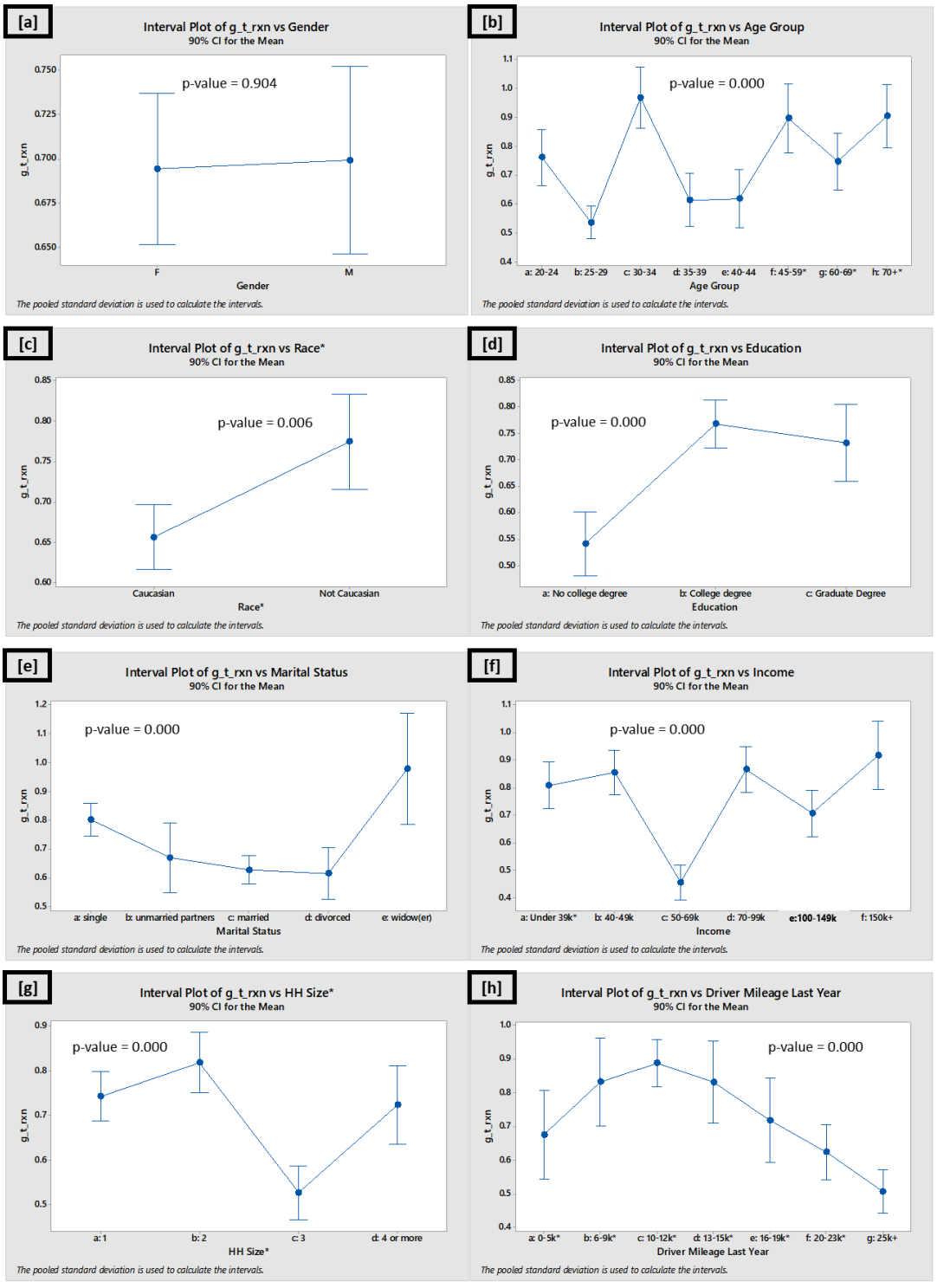


Figure 5.22 Estimated Gipps Reaction Time (t_rxn) Coefficient Segmented by Driver Attributes

As illustrated in Figure 5.23 and Table A.23, the smallest average desired velocity coefficient was estimated for the widow(er) subcategory (mean = 28.1 m/s). The maximum average desired velocity coefficient was estimated for the unmarried partners subcategory (mean = 35.0 m/s). It should be noted that these trends are consistent with the desired velocity parameter estimated for the W99 car-following model, documented in Figure 5.17 and Table A.17. The smallest within subcategory variation was observed for the divorced subcategory (standard deviation = 1.7 m/s). The largest within subcategory variation was calculated for the 60–69 age group (standard deviation = 4.4 m/s). The smallest within attribute variation was calculated for the age attribute [mean(standard deviation) = 3.1 m/s], while the largest within attribute variation was calculated for the household size attribute [mean(standard deviation) = 3.6 m/s].

Figure 5.24 illustrates how the IDM maximum desired acceleration (a) estimated parameter coefficient varies across driver specific attributes. The maximum desired acceleration estimated coefficient varied across all subcategories of driver attributes except gender and age.

According to Table A.24, the smallest average maximum desired acceleration coefficient was estimated for the widow(er) subcategory of the marital status attribute (mean = 0.32 m/s^2); this subcategory was also observed to have the smallest within subcategory variation (standard deviation = 0.32 m/s^2). The largest average maximum desired acceleration parameter coefficient was estimated for the $40–49k income category (mean = 1.39 m/s^2). The largest within subcategory variation was observed to occur within the 35–39 age group (standard deviation = 1.06 m/s^2). The smallest average within attribute variation across the subcategories was observed for the marital status attribute [mean(standard deviation) = 0.76 m/s^2]. The largest average within attribute variation was calculated for the household size attribute [mean(standard deviation) = 0.84 m/s^2).

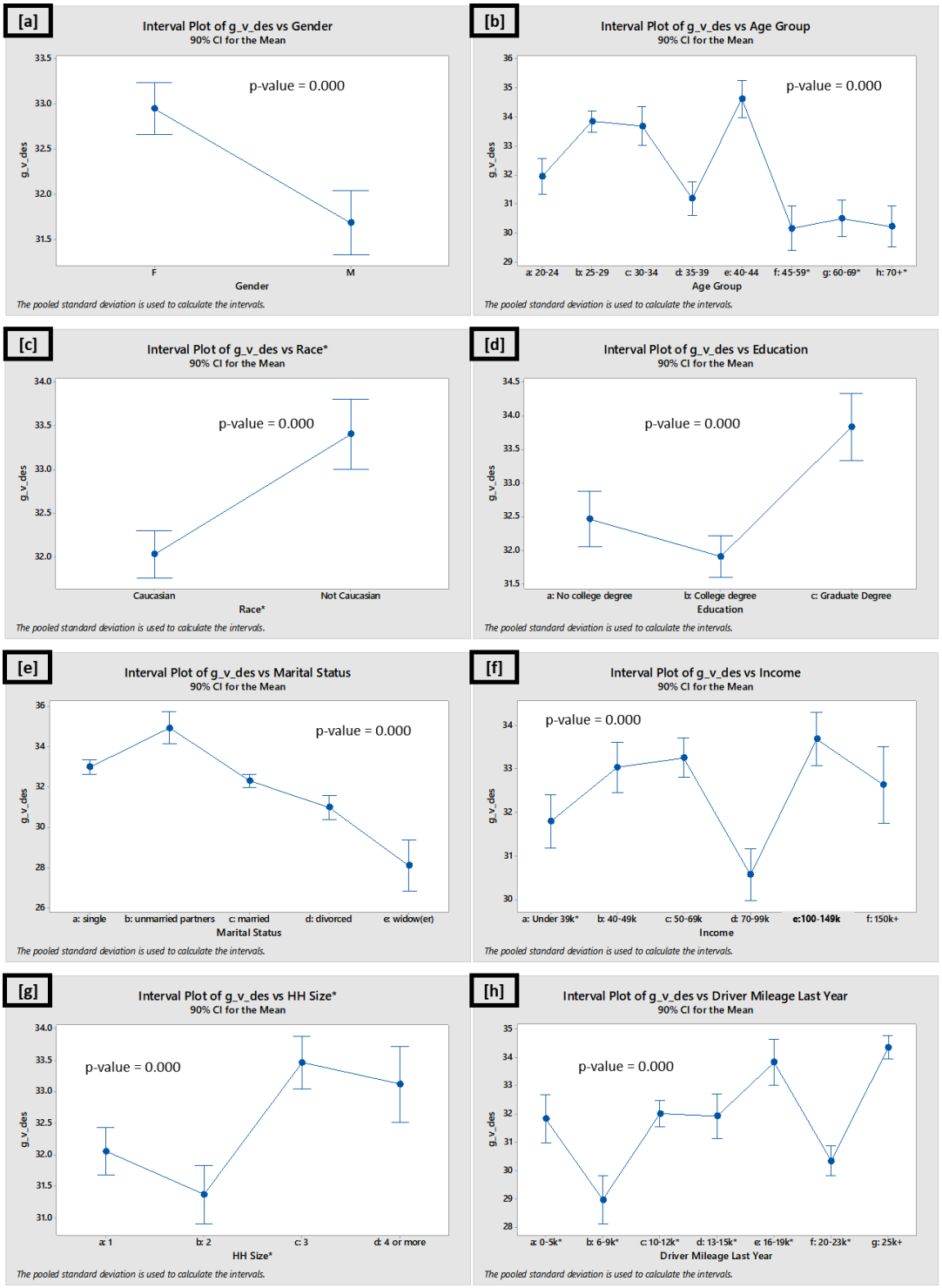


Figure 5.23 Estimated Gipps Desired Velocity (v_des) Coefficient Segmented by Driver Attributes

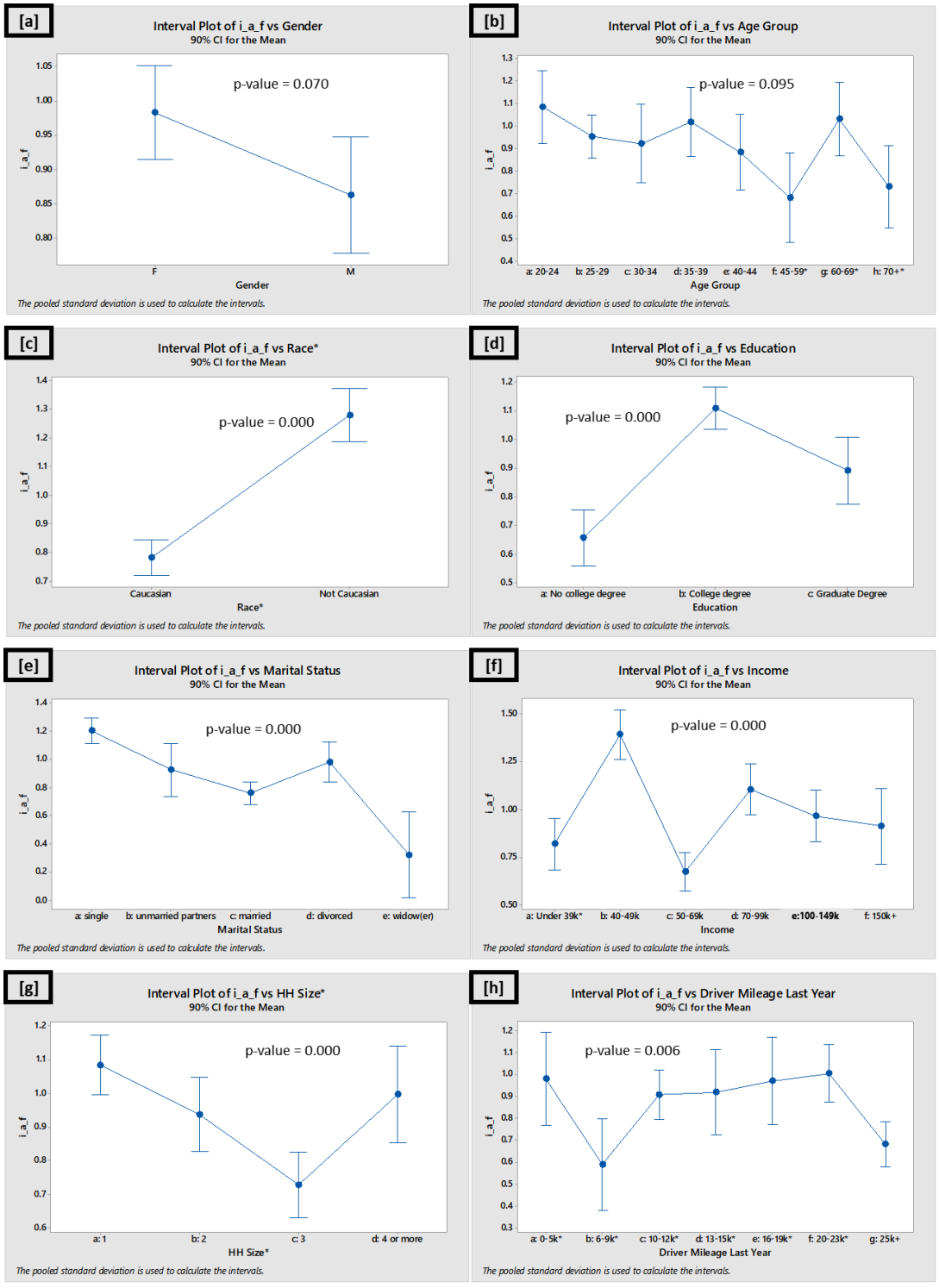


Figure 5.24 Estimated Intelligent Driver Model Maximum Desired Acceleration (a) Coefficient Segmented by Driver Attributes

Figure 5.25 portrays how the IDM maximum desired deceleration (b) estimated coefficient varies across driver specific attributes. The maximum desired deceleration parameter estimates were found to vary at a statistically significant level for all subcategories of driver attributes. According to Figure 5.25a, the average maximum desired acceleration coefficient for male drivers is larger than that estimated for female drivers. In Figure 5.25b, it is observed that the estimated maximum desired deceleration rate for middle age drivers is higher than for younger and older drivers; this observation is consistent with trends observed across the calibrated estimates for the Gipps maximum desired deceleration parameter in Figure 5.19b. Figure 5.25c suggests that the estimated maximum desired deceleration rate is higher for Caucasian driver than it is for drivers who do not identify as Caucasian. Lastly, Figure 5.25g indicates that household size and estimated maximum desired deceleration rate are positively correlated: as household size increases, maximum desired deceleration rate increases.

According to Table A.25, the smallest average maximum desired deceleration rate was estimated for the \$40–49k subcategory of the income attribute (mean = 1.4 m/s^2). The largest average maximum desired deceleration rate was observed for the unmarried partners subcategory of the marital status attribute (mean = 3.2 m/s^2); this attribute was also observed to have the smallest within subcategory variation (standard deviation = 1.07 m/s^2). The subcategory with the largest variation was the widow(er) subcategory (standard deviation = 1.60 m/s^2). The marital status attribute was observed to have the smallest average within attribute variation [mean(standard deviation) = 1.35 m/s^2]; the household size attribute was calculated to have the largest [mean(standard deviation) = 1.41 m/s^2].

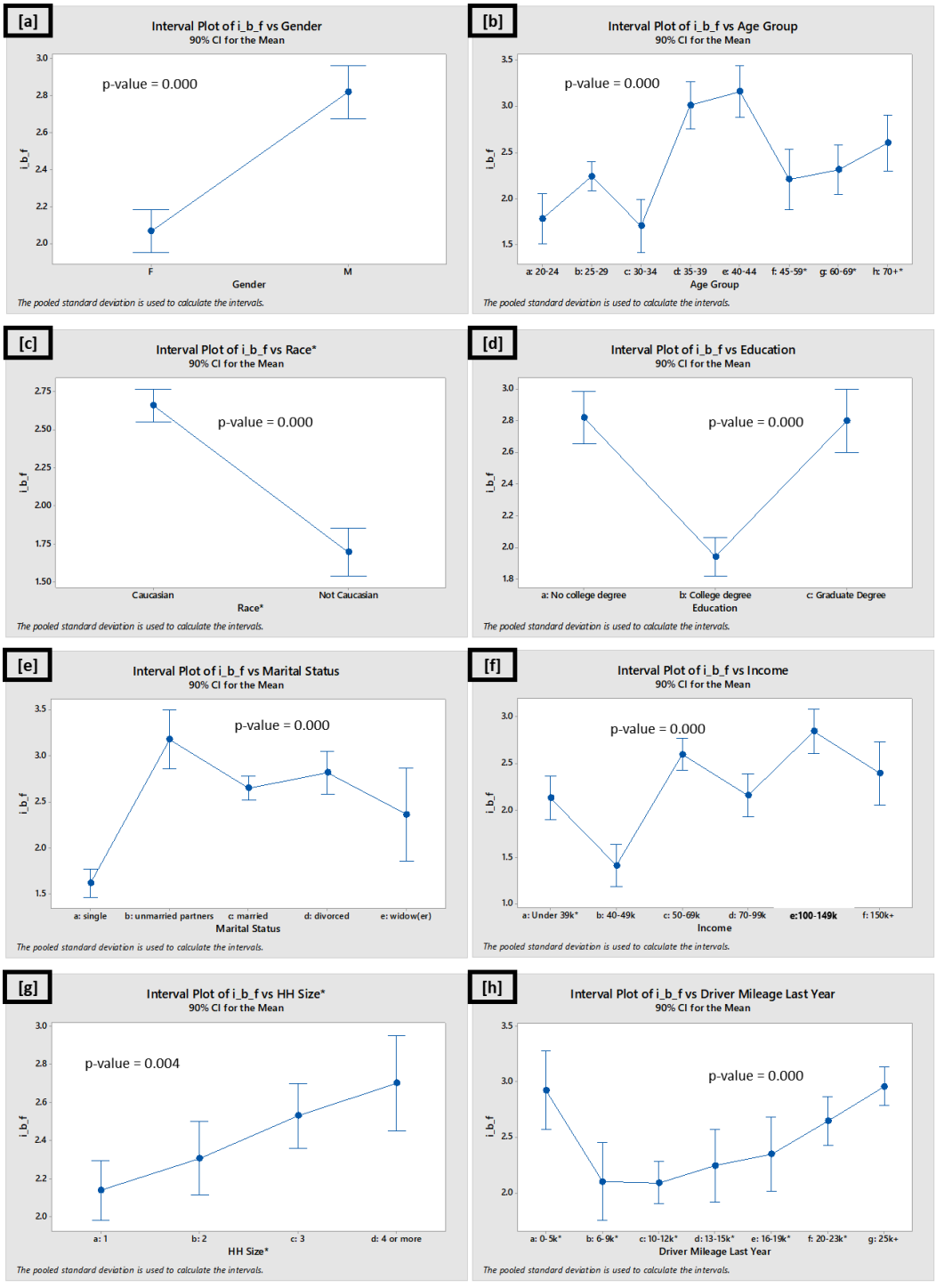


Figure 5.25 Estimated Intelligent Driver Model Maximum Desired Deceleration (b) Coefficient Segmented by Driver Attributes

Figure 5.26 shows how the IDM free acceleration exponent (δ) estimated parameter coefficient varies across driver specific attributes. This parameter describes how a driver's acceleration decreases as the velocity approaches the desired velocity. The free acceleration exponent was found vary at a statistically significant level across all attributes except gender and age.

According to Table A.26, the smallest average δ parameter coefficient was estimated for the widow(er) subcategory (mean = 24.6); the widow(er) subcategory also achieved the smallest within subcategory variation (standard deviation = 28.6). The largest average δ parameter coefficient was estimated for the divorced subcategory (mean = 59.4). The largest within subcategory variation was observed for the 60–69 age group (standard deviation = 41.7). The smallest average variation across an attribute was observed for marital status [mean(standard deviation) = 35.8]. The largest average variation across an attribute was observed for driver age [mean(standard deviation) = 38.6].

Figure 5.27 portrays how the IDM jam distance (g_min) estimated coefficient varies across driver attributes. The jam distance estimated coefficient varied at a statistically significant level for half of the driver attributes: age, race, marital status, and income. According to Table A.27, the largest average parameter estimate for jam distance was found for the unmarried partner subcategory (mean = 5.3 m). The smallest average jam distance coefficient was estimated for the 6–9k miles driven last year subcategory (mean = 3.2 m). The smallest within subcategory variation was calculated for the $40–49k income subcategory (standard deviation = 2.3 m). The largest within subcategory variation occurred in the widow(er) subcategory (standard deviation = 3.9 m). The smallest average within attribute variation was calculated for the race attribute [mean(standard deviation) = 3.1 m], while the largest average within attribute variation occurred in the age attribute [mean(standard deviation) = 3.3 m].

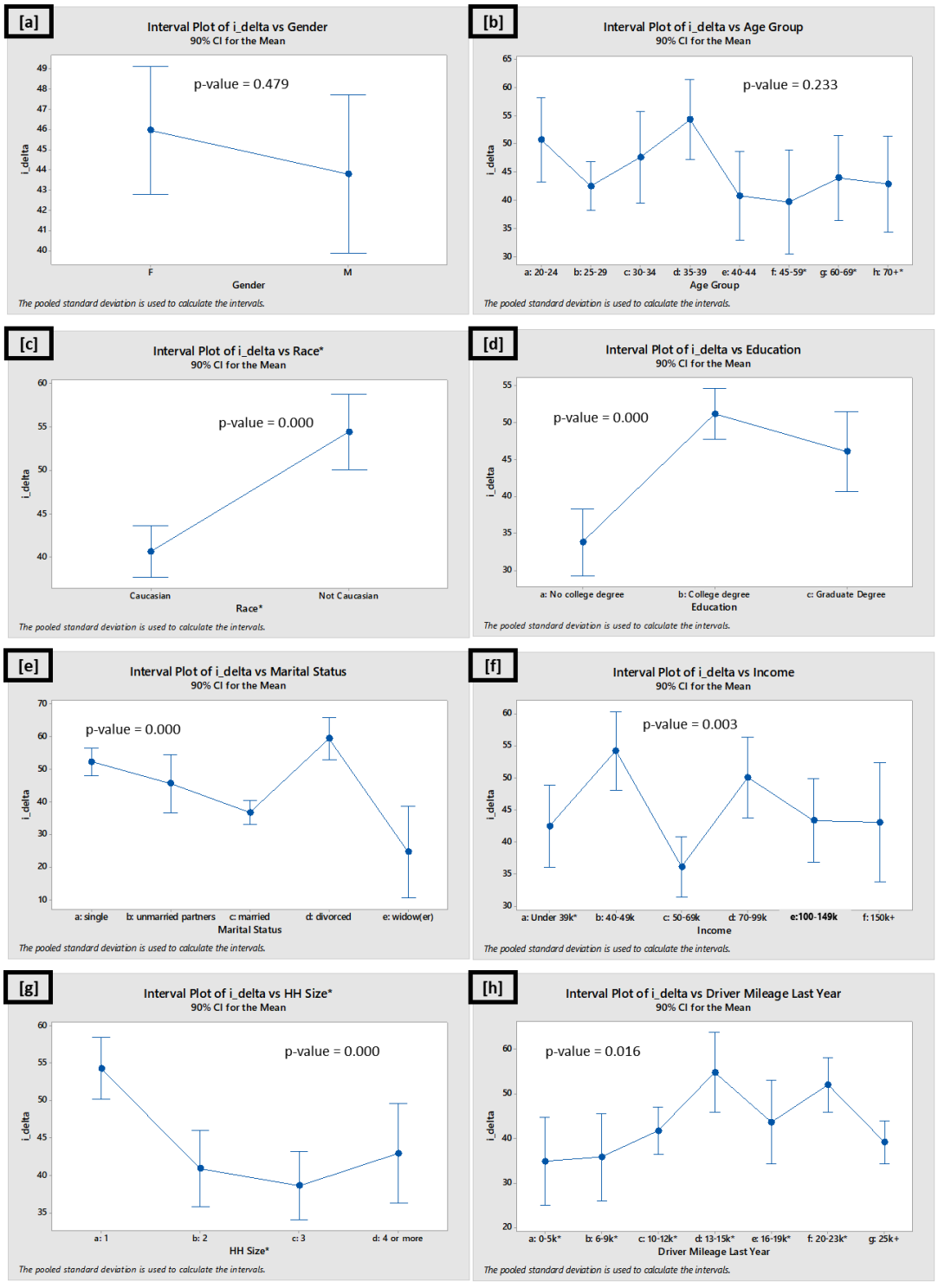


Figure 5.26 Estimated Intelligent Driver Model Free Acceleration Component (δ) Coefficient Segmented by Driver Attributes

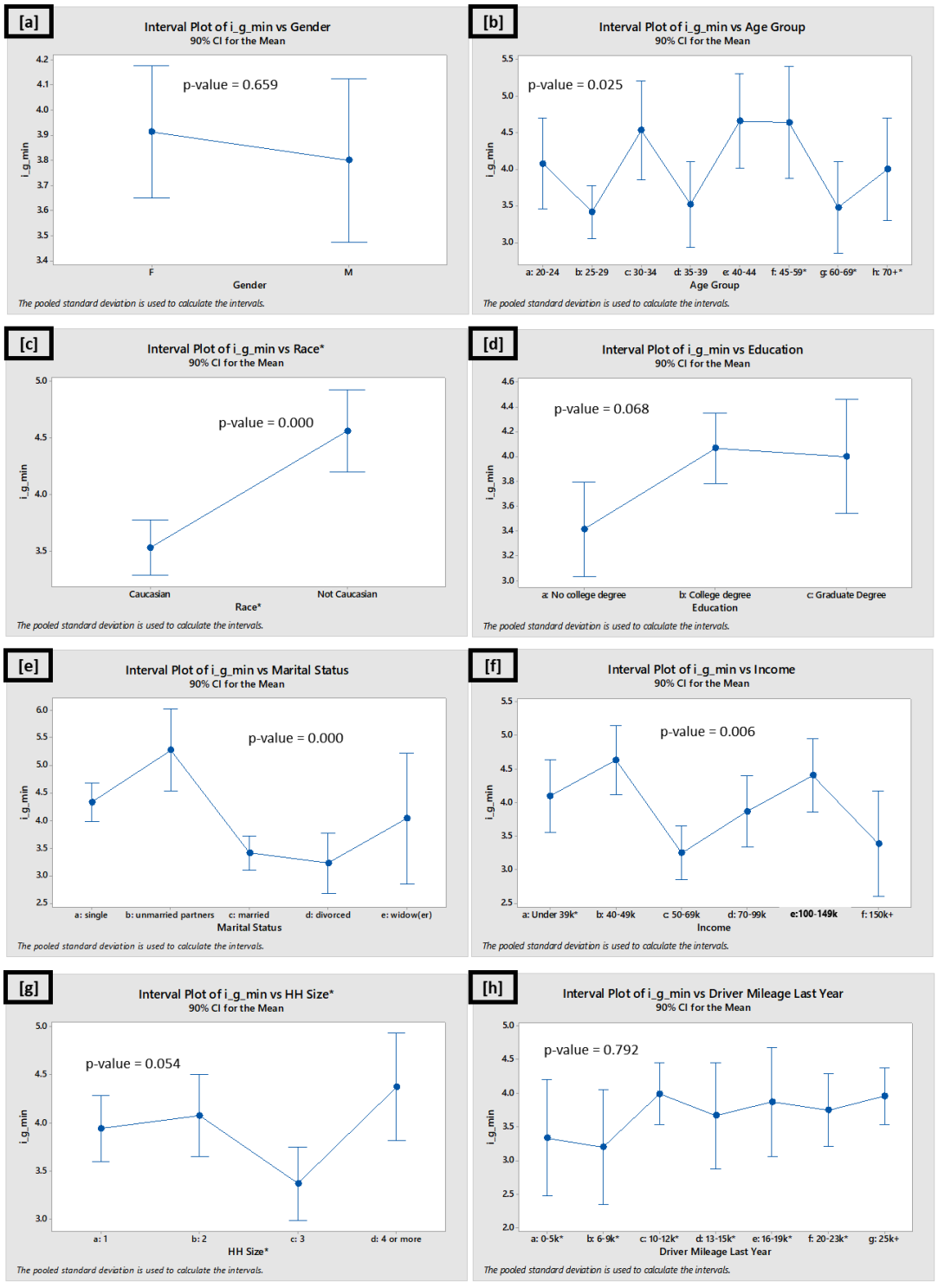


Figure 5.27 Estimated Intelligent Driver Model Jam Distance (g_min) Coefficient Segmented by Driver Attributes

Figure 5.28 illustrates how the IDM desired time gap (t_gap) estimated coefficient varies according to driver attributes. The estimated desired time gap coefficient varies at a statistically significant level across all subcategories of driver attributes, excluding race.

According to Table A.28, the smallest average desired time gap coefficient was estimated for the 25–29 age group (mean = 0.47 s); this subcategory also achieved the smallest within subcategory variation (standard deviation = 0.20 s). The largest average desired time gap coefficient was estimated for the widow(er) subcategory (mean = 1.73 s); this subcategory was also found to have the largest within subcategory variation (standard deviation = 1.52 s). The smallest average attribute variation across subcategories was calculated for the household size attribute [mean(standard deviation) = 0.50 s]; conversely, the largest average attribute variation was calculated for the driver mileage last year attribute [mean(standard deviation) = 0.55 s].

Figure 5.29 documents how the IDM desired velocity attribute estimated coefficient changes across driver attributes. The desired velocity coefficient estimates were found to vary at a statistically significant level across all subcategories of driver attributes. As illustrated in Figure 5.29a, the estimated desired velocity of male drivers is lower than that estimated for female drivers. Figure 5.29b suggests that desired speed and driver age are inversely correlated; the 40–44 age group does not fit this trend, with a significant spike in estimated value compared to the adjacent subcategories. Figure 5.29f may suggest that desired speed is somewhat positively correlated with driver income; the $70–99k subcategory does not fit this trend, with a significant drop in estimated value compared to adjacent categories. Figure 5.29e illustrates that marital status and desired velocity are correlated.

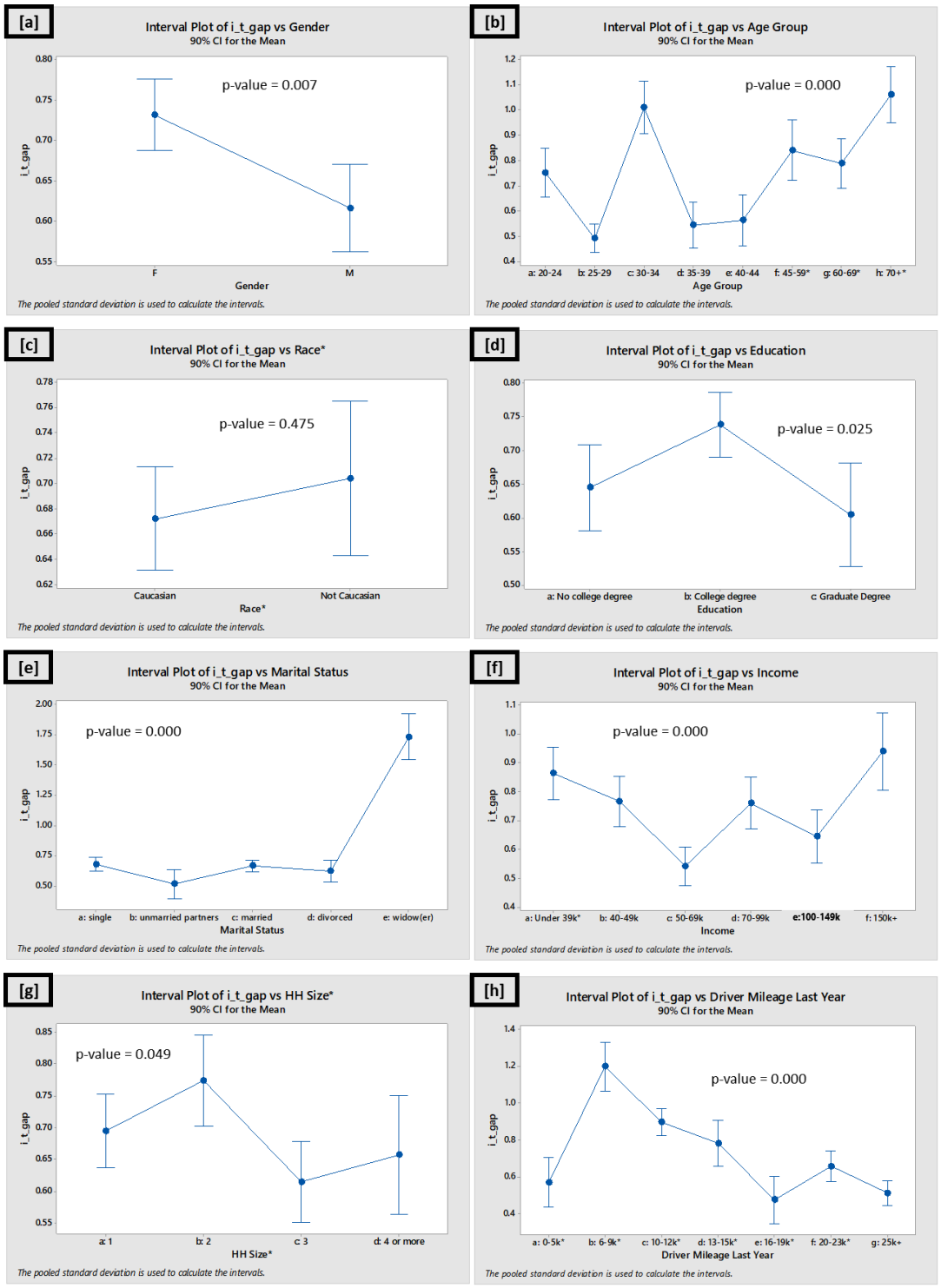


Figure 5.28 Estimated Intelligent Driver Model Desired Time Gap (t_gap) Coefficient Segmented by Driver Attributes

The smallest desired velocity coefficient was estimated for the widow(er) subcategory (mean = 29.4 m/s). The unmarried partner subcategory was estimated to have the largest average desired velocity coefficient (35.8 m/s). The smallest within subcategory variation was calculated for the divorced subcategory of the marital status attribute (standard deviation = 2.1 m/s). The largest within subcategory variation was calculated for the 13–15k miles driven last year subcategory (standard deviation = 5.3 m/s). The smallest average within attribute variation across subcategories was calculated for the age attribute [mean(standard deviation) = 3.1 m/s]. The largest within attribute variation was found in the educational attainment attribute [mean(standard deviation) = 3.5 m/s].

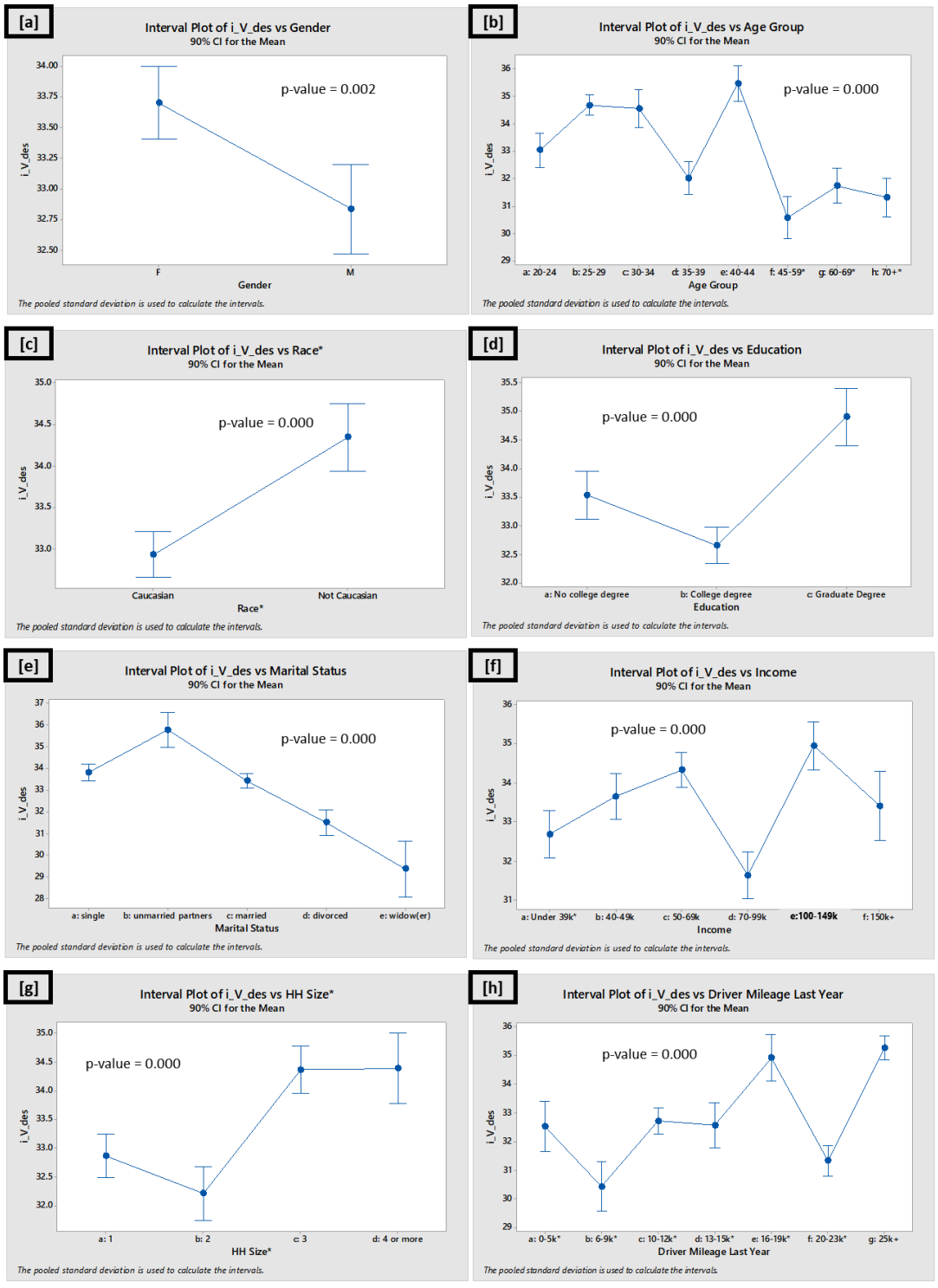


Figure 5.29 Estimated Intelligent Driver Model Desired Velocity (v_des) Coefficient Segmented by Driver Attributes

5.2.2.3. Comparing Results across Car-Following Models: Shared Parameters

In the previous section, this dissertation evaluates trends in calibrated parameter estimates for three car-following models across eight available driver attributes in the SHRP2 NDS. These three car-following models are comprised of a total of 29 different calibration parameters. Although the three car-following models have different functional forms, several calibration parameters have similar physical interpretations across the models. Therefore, it is an interesting exercise to see how the parameters with similar physical meanings relate across the different model structures. In addition, consistent trends and estimates for these similarly defined parameters helps to increase confidence in the model calibration procedure, which was solved with a heuristic instead of an exact solution method. There are five physically interpretable parameters that are seen in two or more of the calibrated models:

- Time gap: (i) W99 spacing time (cc1) and (ii) IDM desired time gap (t_gap)
- Maximum acceleration rate: (i) Gipps maximum desired acceleration rate (a_des) and (ii) IDM maximum desired acceleration rate (a);
- Maximum deceleration rate: (i) Gipps maximum desired deceleration rate (d_des) and (ii) IDM maximum desired deceleration rate (b);
- Inter-vehicle spacing at a stop: (i) W99 standstill distance (cc0), (ii) Gipps minimum gap at a stop (g_min), and (iii) IDM jam distance (g_min); and
- Desired velocity: (i) W99 desired velocity (v_des), (ii) Gipps desired velocity (v_des), and (iii) IDM desired velocity (v_des).

The W99 spacing time and the IDM desired time gap estimated coefficients are documented in Figure 5.8 and Figure 5.28, respectively. There were more driver attributes exhibiting statistically significant differences in estimated parameter coefficients for the IDM desired time gap parameter; however, race was not found to vary at a statistically

significant level for either parameter. The 25–29 age group was estimated to have the smallest value for both parameters, with average coefficients of 0.58 s and 0.47 s, respectively. Moreover, the widow(er) subcategory was estimated to have the highest parameter coefficients, with average values of 2.16 s and 1.73 s, respectively. For both parameters, the 25–29 age group and widow(er) subcategories were found to have the smallest and largest within subcategory variation, respectively.

The Gipps and IDM maximum desired acceleration estimated parameter coefficients are documented in Figure 5.18 and Figure 5.24, respectively. The $40–49k income subcategory was estimated to have the largest average maximum desired acceleration rate for both parameters, with values of 1.8 m/s^2 and 1.4 m/s^2, respectively. Moreover, the widow(er) subcategory was estimated to have the smallest average maximum desired acceleration rate for both models, with estimates of 0.6 m/s^2 and 0.3 m/s^2, respectively; this subcategory achieved the smallest within subcategory variation for both estimated parameters. Additionally, the marital status attribute was calculated to have the smallest within attribute variation for both estimated parameters.

The Gipps and IDM maximum desired deceleration estimated parameter coefficients are documented in Figure 5.19 and Figure 5.25, respectively. Numerous similarities in trends are documented. All subcategories comprising driver specific attributes were found to vary at statistically significant levels for both calibration parameters. Male drivers were estimated to have larger maximum desired deceleration parameter coefficients than female drivers for both models. Additionally, both models were estimated to have larger maximum desired deceleration rates for middle age drivers than for younger or older driver groups. Furthermore, household size and maximum desired deceleration rate were found to be positively correlated for both models. Although the subcategories documented as having the smallest maximum desired deceleration rate were

slightly different (widow(er) and $40–49k income subcategory, respectively), the unmarried partners subcategory was assigned the largest average parameter estimate for both models, with average coefficients of 3.6 m/s^2 and 3.18 m/s^2, respectively.

Parameters loosely representing the minimum inter-vehicle spacing when vehicles are stopped are included in all three models: W99 standstill distance (Figure 5.7), Gipps minimum distance at a stop (Figure 5.21), and IDM jam distance (Figure 5.27). Of the five examples of similarly interpretable model parameters, this attribute is documented as having the least similarities across the different models. The estimated parameter coefficients were not found to vary at a statistically significant level across subcategories for the driver gender or reported mileage driven last year attributes for any of the three parameters. Additionally, across all three models, the race attribute achieved the lowest average within attribute variation, while the mileage driven last year attribute achieved the highest average variation for the Gipps and W99 parameters.

Finally, the W99, Gipps, and IDM models each have a desired velocity estimated parameter; these estimates are included in Figure 5.17, Figure 5.23, and Figure 5.29. All subcategories of driver attributes were found to vary at statistically significant levels for each of the three estimated parameters. Across all three models, male drivers were estimated to have lower desired velocities than female drivers. Additionally, the results suggest that driver age and desired velocity are inversely correlated; for each of the three models, the 40–44 age group appears to be an exception, as this subcategory has a substantially higher estimated coefficient than the adjacent subcategories. The results also suggest that desired velocity and income are somewhat correlated for each of the three models, with an exception for the $70–99k income group; this income group has a considerably lower average estimated parameter coefficient than adjacent categories across each of the three model parameters. The widow(er) subcategory was estimated to have the

smallest average desired velocity coefficient for each of the three models, with desired speeds of 27.3 m/s, 28.1 m/s, and 29.4 m/s, respectively. Additionally, the unmarried partners subcategory was estimated to have the largest average desired velocity values across all three models: 35.2 m/s, 34.9 m/s, and 35.8 m/s, respectively.

5.2.3. Conclusions

In conclusion, this effort explores research question number four posed by this dissertation: when subsamples of drivers are created using driver attributes, are the parameter sets for the different subsamples of drivers sufficiently different to justify the segmentation of the data?

To answer this question, three widely accepted car-following models are calibrated to a 665-trip sample of the SHRP2 NDS dataset. Eight driver attributes collected in questionnaires as part of the SHRP2 NDS data collection effort—gender, age, race, educational attainment, marital status, income, household size, and self-reported driver mileage last year—are used to segment data into hypothesized subgroups of data.

The purpose of this chapter is to provide support for Chapters 8 and 9 of this dissertation. Toward this end, this chapter provides significant evidence that there are statistically significant differences in driver behavior (i.e., calibrated car-following model parameter estimates) between different subcategories (e.g., male, female) of driver attributes (e.g., gender). Segmenting the data across some attributes, such as driver age and marital status, produced more conclusive trends than other attributes, such as gender and driver mileage last year.

5.3. Similarities in Driving Behavior Within Subgroups of Drivers Segmented by Driver Specific Attributes

The first objective of this research is to identify homogeneous groups of driving behavior. As mentioned throughout this dissertation, there is substantial evidence of variation in driving behavior in naturalistic data (A. L. Berthaume et al., 2018; James, Hammit, & Ahmed, 2018; Ossen & Hoogendoorn, 2011). This research seeks to determine if homogeneous groups of driving behavior exist within the data: that is, do groups of drivers that behave sufficiently similarly to one another for inclusion in a single homogeneous group exist within the data?

Should question 1 be true, it is important to understand the similarities between drivers clustered into the same group of homogeneous driving behavior (i.e., if Driver A and Driver B are placed into the same group, what other commonalities do they have beyond those used in the similarity assessment?). It is equally important to understand the differences between groups of homogeneous drivers (i.e., if Driver A and B are in group 1, while Driver C and D are in group 2, how are Driver A/B and C/D different?). This dissertation hypothesizes that driver attributes can be used to identify commonalities within a group and differences between groups of homogeneous driving behavior.

5.3.1. Methodology

To study these two research questions, first, the Gipps, IDM, and W99 car-following models are calibrated independently for each of the 665 trips in the available sample of SHRP2 NDS data; this procedure is detailed in Section 3.2.4. This produces almost 2000 trip-specific sets of calibration coefficients for analysis.

Clustering methods were identified as appropriate tools to explore question 1; the goal of a clustering analysis is to identify a specified number of classes within the data such

that an instance placed in a class is sufficiently similar to the other instances placed within the class and sufficiently different to the instances in other classes. Specifically, the Expectation Maximization algorithm within Weka, or the Waikato Environment for Knowledge Analysis, was selected as the clustering algorithm for this analysis. The Expectation Maximization algorithm is a soft clustering algorithm; it assigns instances to classes probabilistically and is capable of identifying the optimal number of clusters. The Expectation Maximization algorithm has similar intuition to the k-means clustering algorithm, but does not produce hard clusters. For k-means, when a cluster is placed, only the points within the cluster boundary belong and contribute to the calculation of the cluster centroid. For Expectation Maximization, the clusters can be thought of as a distribution and boundaries are less rigid. Every single data point will contribute to the location of cluster centroids in Expectation Maximization, with data points strongly associated with a cluster providing more influence to that cluster's centroid than other data points (Witten, Frank, Hall, & Pal, 2017d).

Assigning cluster IDs probabilistically requires an iterative solution: what comes first, the apple or the egg? If one knew which instance belonged to which distribution or if one knew the mean and variance of the normal distributions to which the instances belonged, assigning cluster IDs would be trivial. Given that that information is not initially available the Expectation Maximization algorithm is required.

To initialize the Expectation Maximization algorithm, *k* Gaussian distributions (i.e., clusters) are placed with an assumed mean (μ) and variance (σ^2). For conciseness, this explanation will assume *k=2* (Cluster A and Cluster B), though this algorithm extends beyond two clusters.

The mean and variance parameters of each Gaussian are used to calculate the probability that each data point is associated with each distribution. Equations 5.1 and 5.2 determine if an instance appears more like a sample from Cluster A or Cluster B:

$$P(x_i \mid a) = \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left(-\frac{(x_i - \mu_a)^2}{2\sigma_a^2}\right) \quad 5.1$$

$$P(x_i \mid b) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left(-\frac{(x_i - \mu_b)^2}{2\sigma_b^2}\right) \quad 5.2$$

Next, the points are softly assigned to the distributions through posterior probabilities, which are conditional probabilities assigned after relevant evidence (i.e., means and variances of the distributions) is accounted for. Equations 5.3 and 5.4 determine the degree to which each point belongs to Cluster A and Cluster B (i.e., soft clustering):

$$a_i = P(a_i|x_i) = \frac{P(x_i \mid a)\, P(a)}{P(x_i \mid a)P(a) + P(x_i \mid b)P(b)} \quad 5.3$$

$$b_i = P(b_i|x_i) = 1 - a_i \quad 5.4$$

Next, the means and variances of the distributions are adjusted based on the probability of the points belonging to each of the distributions:

$$\mu_a = \frac{a_1x_1 + a_2x_2 + \cdots + a_nx_n}{a_1 + a_2 + \cdots + a_n} \quad 5.5$$

$$\sigma_a^2 = \frac{a_1(x_1 - \mu_a) + \cdots + a_n(x_n - \mu_a)}{a_1 + \cdots + a_n} \quad 5.6$$

$$\mu_b = \frac{bx_1 + b_2x_2 + \cdots + b_nx_n}{b_1 + b_2 + \cdots + b_n} \quad 5.7$$

$$\sigma_b^2 = \frac{b_1(x_1 - \mu_b) + \cdots + b_n(x_n - \mu_b)}{b_1 + \cdots + b_n} \quad 5.8$$

This procedure—identifying the probability that each point belongs to each cluster and re-estimating the parameters of the Gaussian distributions for each cluster based on the probabilities—repeats until convergence.

An example of the results of the Expectation Maximization algorithm for one of the parameters is shown in Figure 5.30. The Expectation Maximization algorithm assigns each trip (i.e., driver) a cluster ID, indicating to which cluster of homogeneous car-following model parameter coefficients driver belongs.

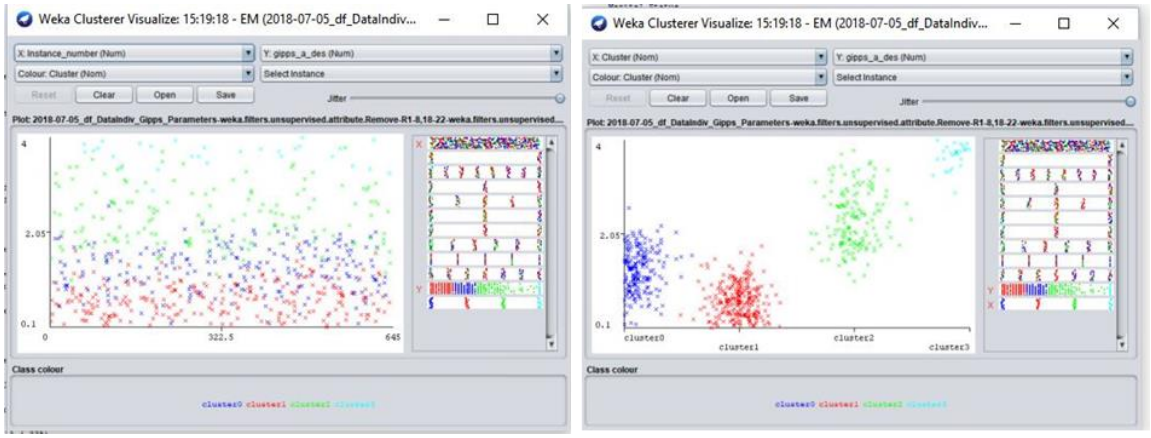


Figure 5.30 Clusters Identified by the Expectations Maximization Algorithm for the Maximum Desired Acceleration Parameter of the Gipps Model (James & Hammit, 2019)

Classification methods were explored as tools to provide insight into question 2. This effort classifies the data based on similarity/differences in driver attributes (e.g., age, gender, marital status). Four classification methods were utilized in this dissertation:

- *ZeroR*: Predicts that all data belongs to the majority class (Witten, Frank, Hall, & Pal, 2017b). For the purpose of this paper, this is the "baseline" accuracy measure (i.e., model performance without using driver attributes to improve the classification results). This is analogous to a modeler using the most common driving behavior to describe all of the observed behavior.

- *OneR*: Produces a one-level decision tree in the form of a set of rules for the most predictive attribute (Witten et al., 2017b).
- Divide-and-conquer strategies: The goal of divide-and-conquer strategies is to maximize the separation between classes by choosing attributes that maximize the information gain on which to split. The *J48* Decision Tree and *PART* Decision Rules algorithms are both types of divide-and-conquer classification strategies (Witten et al., 2017b)
    - *J48* Decision Tree: Recursively builds a C4.5 decision tree, taking all classes into account when making decisions. The logic of the *J48* algorithm is as follows: (i) select attributes for root node, creating one branch for each attribute value; (ii) split instances into subsets, one for each branch extending from the node; and (iii) recursively repeating the process for each branch, only using attributes that reach that branch (Witten et al., 2017b).
    - *PART* Decision Rules: Derives classification rules using a C4.5 inspired heuristic, concentrating on one class at a time. The logic of the *PART* algorithm is as follows: (i) make a rule by building a tree and reading off the rule for the largest leaf; (ii) remove all instances this rule covers; (iii) discard the tree; and (iv) repeat steps for remaining data not covered by previously developed rules (Witten et al., 2017b).

The success of the classification method is determined by the model's ability to accurately assign a trip to the cluster it belongs to only using attributes about the driver; this is referred to as the model's accuracy rate, or the number of trips correctly assigned to

the homogeneous behavioral cluster to which it belongs. A 10-fold cross-validation protocol was used to robustly estimate the accuracy rate of the classification algorithms.

Attribute selection methods were used to help eliminate redundant and irrelevant attributes before the classification procedures were applied (Witten, Frank, Hall, & Pal, 2017c). Specifically, the *CfsSubsetEval* method in Weka was utilized, which uses an entropy-based metric called symmetric uncertainty to explore the predictive power and the degree of inter-redundancy of each attribute before making independent variable recommendations in the classification problem.

The framework used to cluster and classify driving behavior by driver attributes is shown in Figure 5.31. The process is repeated independently for each of the calibration model parameters from each car-following model.

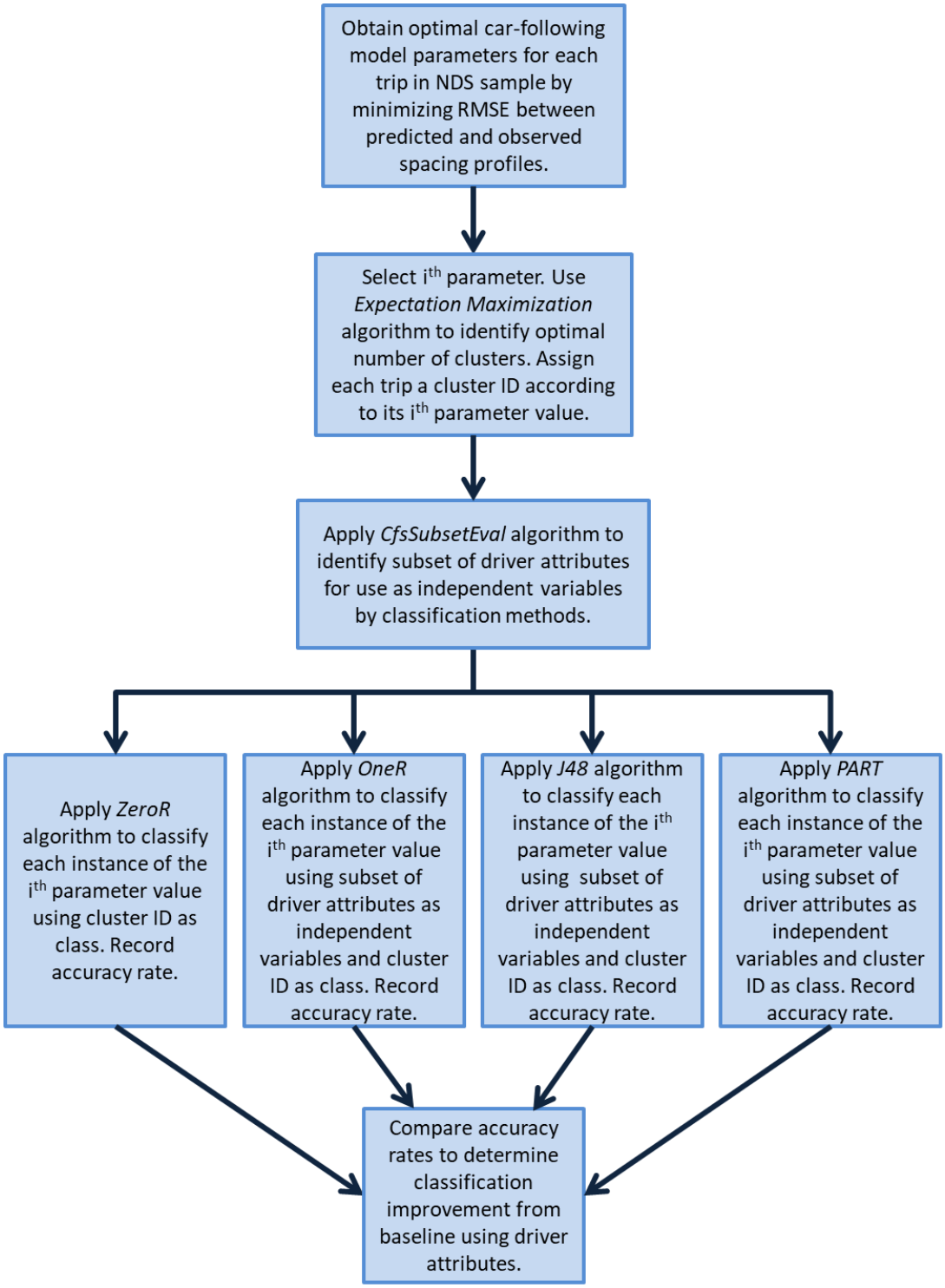


Figure 5.31 Framework of Proposed Procedure to Cluster and Classify Calibrated Parameter Coefficients (James & Hammit, 2019)

5.3.2. Results

The results of the attribute selection process for the W99, Gipps, and IDM, models are available in Table 5.1, Table 5.2, and Table 5.3, respectively. The results of the clustering and classification methods for the W99, Gipps, and IDM models are available in Table 5.4, Table 5.5, and Table 5.6, respectively.

5.3.2.1. Attribute Selection

As evident in Table 5.1, Table 5.2, and Table 5.3, the most important driver attributes vary between each parameter. The most commonly selected attributes were age (21 parameters), marital status (18 parameters), and income (11 parameters). The least frequently selected attributes were gender, race, and educational attainment. This is somewhat corroborative with the literature review, where age was more commonly used than gender to reduce heterogeneity in desired speed and headway data (Boyce & Geller, 2002; Elander et al., 1993). The frequency of the selection of marital status as a critical attribute in the identification of clustered behavior was a bit surprising; the authors postulate that one possible explanation for this observation could be the presence of passengers (e.g., spouse, children) in the vehicle, but this is difficult to verify given the data available.

Table 5.1 W99 Parameters Attribute Selection Results

Model Calibration Parameters	***CfsSubsetEval* Results**
Standstill Distance (cc0)	Marital Status, Income
Spacing Time (cc1)	Age Group, Driver Mileage Last Year
Following Variation, Maximum Drift (cc2)	Gender, Age Group, Marital Status
Threshold for Entering 'Following' (cc3)	Age Group, Marital Status, Income, Driver Mileage Last Year
Negative Following Threshold (cc4)	Age Group, Marital Status, Household Size
Positive Following Threshold (cc5)	Age Group, Marital Status, Household Size
Speed Dependency of Oscillation (cc6)	Gender, Age Group, Marital Status, Driver Mileage
Oscillation Acceleration (cc7)	Age Group, Marital Status, Race
Standstill Acceleration (cc8)	Age Group, Race, Income, Household Size
Acceleration at 80 kph (cc9)	Age Group, Marital Status, Driver Mileage
Desired Speed (vdes)	Age Group, Marital Status, Driver Mileage Last Year
All W99 Parameters	Age Group, Marital Status

Table 5.2 Gipps Parameters Attribute Selection Results

Model Calibration Parameter	***CfsSubsetEval* Results**
Maximum Desired Acceleration (a)	Age Group, Education, Marital Status, Living Status, Work Status, Income
Maximum Desired Deceleration (b)	Gender, Age Group, Education, Marital Status, Income
Perception of Leading Vehicle's Desired Braking ($\hat{b}$)	Age Group, Education, Marital Status, Income, Driver Mileage Last Year
Minimum Following Distance at a Stop (s_0)	Marital Status, Age Group, Income
Reaction Time (τ)	Age Group, Income, Driver Mileage Last Year
Desired Velocity (V)	Age Group, Marital Status, Driver Mileage Last Year
All Gipps Parameters	Age Group, Marital Status, Education, Income, Household Size, Gender

Table 5.3 IDM Parameters Attribute Selection Results

Model Calibration Parameter	***CfsSubsetEval* Results**
Maximum Desired Acceleration (a)	Age Group, Household Size
Maximum Desired Deceleration (b)	Gender, Age Group, Education, Marital Status, Income
Free Acceleration Exponent (δ)	Education, Marital Status, Race, Income
Jam Distance (s_0)	Age Group, Marital Status, and Income
Desired Time Gap (T)	Age Group, Driver Mileage Last Year
Desired Velocity (v_0)	Age Group, Marital Status, Driver Mileage Last Year
All IDM Parameters	Marital Status, Age Group, Driver Mileage Last Year

5.3.2.2. Number of Clusters

The clusters are intended to represent pockets of homogeneous driver behavior. The Expectation Maximization algorithm successfully identified clusters of calibrated car-following model parameter coefficients; this provides significant evidence towards the hypothesis that there exist trips (i.e., drivers) that are sufficiently similar to one another to be considered a homogeneous group. The optimal number of clusters varies significantly between each parameter, as shown in Table 5.4, Table 5.5, and Table 5.6. The number of optimal clusters, determined by the *Expectation Maximization* algorithm, varied from three clusters to eleven clusters. Three and four clusters were the most frequently occurring number of optimal clusters, each occurring four times across the parameters.

5.3.2.3. Classification

The classification results are shown in Table 5.4, Table 5.5, and Table 5.6. As discussed in the Methodology section, the ZeroR algorithm produces the accuracy rate

when all data are classified as the majority class; this is considered the baseline classification accuracy rate (i.e., the best one can do without using independent variables in the classification process). The OneR, J48, and PART algorithms were applied to discern if there were underlying relationships between the clusters of homogeneous parameter coefficients and driver specific attributes; this would be evident by an improvement in the accuracy rate when using driver attributes as independent variables.

Table 5.4 W99 Parameters Clustering and Classification Results

Model Calibration Parameter	**Optimal # of Clusters**	***ZeroR* Accuracy Rate**	**Best Model Accuracy Rate**	**Difference From *ZeroR* (%)**	**Best Model**
Standstill Distance (cc0)	6	163	201	23	*OneR*
Spacing Time (cc1)	9	173	216	25	*J48*
Following Variation, Maximum Drift (cc2)	3	450	494	10	*OneR*
Threshold for Entering 'Following' (cc3)	4	345	345	0	*ZeroR*
Negative Following Threshold (cc4)	4	451	454	1	*OneR*
Positive Following Threshold (cc5)	5	214	236	10	*OneR*
Speed Dependency of Oscillation (cc6)	10	114	131	15	*OneR*
Oscillation Acceleration (cc7)	6	264	296	12	*J48*
Standstill Acceleration (cc8)	4	226	229	1	*PART*
Acceleration at 80 kph (cc9)	5	518	518	0	*ZeroR*
Desired Speed (vdes)	4	288	412	43	*PART*
All W99 Parameters	4	196	242	23	*J48*

As shown in Table 5.4, Table 5.5, and Table 5.6, there appears to be a correlation between driver specific attributes and clusters of homogeneous driving behavior. Across all three models, the use of driver attributes to classify drivers by their desired velocity cluster ID was the most successful, with more than a 40% increase in the number of trips correctly classified.

Table 5.5 Gipps Parameters Clustering and Classification Results

Model Calibration Parameter	Optimal # of Clusters	ZeroR Accuracy Rate	Best Model Accuracy Rate	Difference From ZeroR (%)	Best Model
Maximum Desired Acceleration (a)	4	249	299	20	*J48*
Maximum Desired Deceleration (b)	3	312	342	10	*PART*
Perception of Leading Vehicle's Desired Braking ($\hat{b}$)	4	312	330	6	*J48*
Minimum Following Distance at a Stop (s_0)	6	172	182	6	*PART*
Reaction Time (τ)	11	187	184	-2	*ZeroR*
Desired Velocity (V)	5	229	360	57	*J48*
All Gipps Parameters	4	250	298	19%	*PART*

The maximum desired acceleration parameter appears in both the Gipps and IDM parameter lists. For both models, there is an improvement in the accuracy rate of the classification of drivers by their cluster ID of around 25%. The maximum desired deceleration parameter in the Gipps and the IDM results are not quite as consistent. The maximum desired deceleration parameter in the Gipps model sees an increase of 10% in the accuracy rate; for the IDM parameter, the improvement is 45%. This is an example of parameter correlation observed by Kim and Mahmassani (2011) and Punzo et al. (2015), where although the parameter has the same physical interpretation, the calibrated value is biased by its interaction with the other parameters in the model.

Table 5.6 IDM Parameters Clustering and Classification Results

Model Calibration Parameter	**Optimal # of Clusters**	***ZeroR* Accuracy Rate**	**Best Model Accuracy Rate**	**Difference From ZeroR *(%)***	**Best Model**
Maximum Desired Acceleration (a)	6	200	251	26	*PART*
Maximum Desired Deceleration (b)	5	199	288	45	*PART*
Free Acceleration Exponent (δ)	3	291	311	7	*PART*
Jam Distance (s_0)	4	316	311	-2	*PART*
Desired Time Gap (T)	5	249	343	38	*J48*
Desired Velocity (v_0)	5	194	332	71	*J48*
All IDM Parameters	4	185	248	34%	*OneR*

This phenomenon is quite evident in the minimum following distance at a stop parameter, which is roughly included in the Gipps (s_0), IDM (s_0), and the W99 (cc0) car-following models. For the minimum following distance at a stop parameter, the Gipps and W99 model see an improvement in the accuracy rate over the baseline of 6% and 23%, respectively. Conversely, for the IDM, the ZeroR method performs best; the introduction of the driver attributes as explanatory variables causes the classification method to perform more poorly than if they had not been considered at all. The results of the IDM jam distance parameter (s_0) is an example of overfitting, where the model learned on the training data does not generalize to the validation data; the learned model performs worse than the baseline model, as the OneR, J48, and PART algorithms should have all learned a model that assigns all instances to the majority class. Though there are two instances where the use of driver attributes reduces the classification accuracy below that of the baseline model, it seems the benefits of using driver attributes in the classification problem outweigh the risk.

5.3.3. Conclusions

Section 5.3 sought to address two specific questions concerning car-following models and heterogeneity in driving behavior:

1. Can groups of sufficiently similar driving behavior be identified in trajectory-level data?

2. Provided question one proves true, which attributes are valuable in discerning similarities *within* a homogeneous group and differences *between* homogeneous groups?

To explore these questions, a 665 trip sample from the SHRP2 NDS database was obtained and used to calibrate the Gipps, IDM, and W99 car-following models. The calibrated model parameters were used as surrogates to represent driving behavior. The Expectation Maximization clustering algorithm was applied to each of the car-following model parameters to identify the optimal number of clusters of homogeneous groups within the dataset. The successful application of the clustering algorithm provides insight into question 1: there is evidence in naturalistic data that some drivers behave sufficiently similar to one another (and sufficiently different from drivers belonging to a different group) to be considered a homogeneous group of drivers.

Next, attribute selection and classification algorithms were applied to assign drivers to the previously obtained clusters of homogeneous driving behavior purely based on driver attributes; this was intended to inform insights into question 2. Age and marital status were the most commonly selected attributes, while gender, race, and educational attainment were the least commonly selected attributes. Although the improvements varied depending on the model parameter used for clustering, results indicate that clusters of homogeneous car-following model parameters are correlated with driver attributes (i.e., age). Put succinctly, some drivers drive sufficiently alike to form a cluster of similar behavior.

Moreover, driver specific attributes can be carefully utilized to classify drivers into these homogeneous driver groups.

The practical implications of this research are as follows. Previous research has illustrated the importance of modeling an appropriate level of driver heterogeneity, as it impacts the propagation of shockwaves and the projected capacity (Ossen & Hoogendoorn, 2008, 2011). This research illustrates that a modeler may not need to model every single driver differently; by identifying groups of drivers that behave sufficiently similar (e.g., males, young drivers, drivers navigating work zones), modelers may be able to improve the realism of their models, by accounting for heterogeneity, without significantly increasing the complexity. These findings reinforce the need to develop a new framework to represent car-following heterogeneity in microsimulation models, with the ultimate goal of producing more realistic predictions of traffic flow behavior.

5.4. CHAPTER 5 CONCLUSIONS

In this chapter, there were two primary research questions. The first seeks to understand how **different** the driving behaviors are **between** different groups of drivers (e.g., do male drivers exhibit statistically significant differences in driving behaviors than female drivers). The second research question seeks to understand how **similar** the driving behaviors are **within** different groups of drivers (e.g., should females be considered a group of homogeneous drivers, or are the behaviors too inconsistent to be clustered together?).

The first research question was explored in Section 5.2. Section 5.2 provided two fundamental insights critical to this dissertation. The first is that heterogeneity in naturalistic driving behavior does vary according to driver specific attributes. This provides

some of the first evidence that by controlling for different driver attributes, such as age and gender, the heterogeneity in behavior is reduced.

Section 5.3 explores the second research question. Through the successful application of a classification algorithm, this dissertation provides evidence that there does exist groups of drivers that behave sufficiently similar to one another to be considered as a "homogeneous driver group". Moreover, this chapter overwhelmingly provides evidence that driver attributes may be used to successfully determine these homogeneous driver groups, without an *a priori* knowledge of the specific driver's behavior.

Now that it has been established that behavioral heterogeneity is reduced by controlling for driver specific attributes, Chapter 6 of this dissertation seeks to determine if some models are able to capture the behavior of subgroups of drivers, which are segmented by driver attributes, than other models.

Chapter 6: Assessing Relative Model Performance across Driver Attributes (Task 2)

6.1. MOTIVATION

In Chapter 2 of this dissertation several studies were documented observing significant heterogeneity in naturalistic data. In fact, several studies noted that the heterogeneity in driving behavior was so significant that it required different optimal parameter sets and car-following model functional forms to adequately capture this behavioral heterogeneity (Ossen & Hoogendoorn, 2005, 2011). Moreover, adequately capturing this heterogeneity was found to alter microsimulation outputs, including the propagation of shockwaves and overall corridor capacity (Ossen & Hoogendoorn, 2007). Thus, it is a logical extension to wonder if model performance is a function of driver attributes. More specifically, this chapter seeks to answer research question number 3 posed by this dissertation: of common car-following models in the literature, do certain models describe subsamples of drivers better than others (e.g., does the Gipps model describe female drivers better than the Wiedemann 99 (W99) model)?

6.2. METHODOLOGY

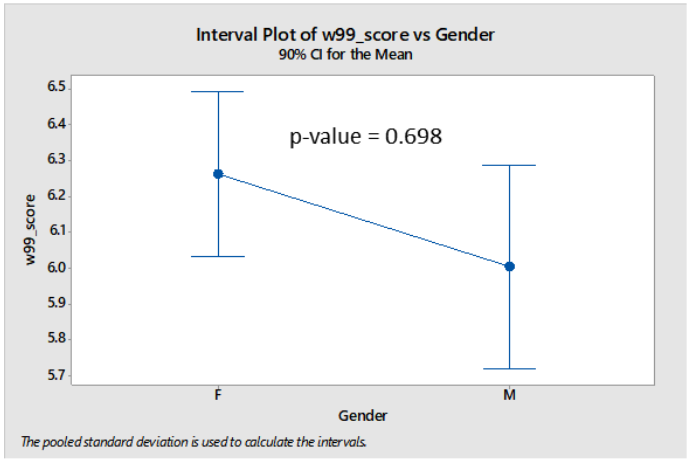
The methodology used to conduct this analysis has three components. The first effort identifies which car-following models to include in the analysis; for details on this effort, the reader is referred to Section 3.2.3. The second effort develops and implements a calibration procedure to identify optimal parameter sets for each trip. Of interest to this chapter is the model calibration score that is attained when the optimal parameter sets are obtained. As a reminder, the calibration objective function seeks to identify the set of calibration parameter coefficients that minimize the root mean square error (RMSE) between the predicted and observed relative spacing profile across the constrained driving

states in a second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) trip; for additional details, the reader is referred to Section 3.2.4. To answer the aforementioned research question, the model calibration scores are split into "subcategories" (e.g., male, female) that comprise a driver "attribute" (e.g., gender). For complete details on the available driver attributes, the reader is referred to Table 3.2. To evaluate if there are significant differences between the subcategories that comprise a driver attribute, the third effort applies statistical tests; this effort is described herein.

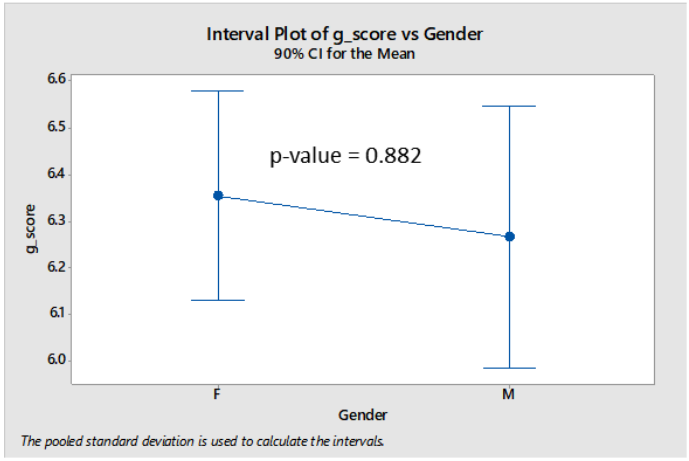
To answer questions regarding the relative performance between models when controlling for driver attributes, Kruskal-Wallis one-way analysis of variance (ANOVA) tests are conducted; this is a non-parametric alterative to the traditional one-way ANOVA test and is required given the non-normality of the distribution of calibration scores, confirmed using Anderson-Darling statistical tests. The Kruskal-Wallis is a type of hypothesis test that seeks to determine if two or more samples are from identical distributions using the distribution medians (i.e., are the distributions statistically different?). This allows one to evaluate the statistical significance of differences in calibration score within a subcategory (e.g., female W99 score vs. female IDM score vs. female Gipps score). The null hypothesis in this test is that the true medians of the population scores are equal (i.e., the model performance for that particular driver subgroup is essentially the same). Thus, small p-values, 5% alpha levels selected for this analysis, result in a rejection of the null hypothesis and the conclusion that the performance of the models at capturing the driving behavior for a subpopulation of drivers is statistically different.

6.3. RESULTS

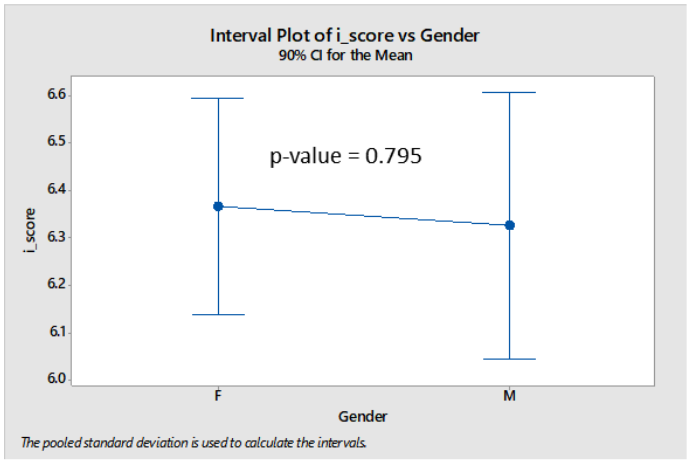
Figure 6.1 illustrates how the W99, Gipps, and Intelligent Driver Model (IDM) calibration scores vary across subcategories of gender. For all three models, the models were able to better predict male driving behavior compared to female behavior, though not at a statistically significant level. The W99 model achieved a lower average and median score compared to the Gipps and IDM car-following models; the differences in male scores across the models were statistically significant, while the differences in female scores were not. However, the W99 model also produced the most variability in scores across a subcategory: the Gipps model achieved the smallest standard deviation in scores across male drivers, while the IDM achieved the smallest standard deviation for female drivers.



W99	Mean	Standard Deviation	Median
Female	6.26*	2.75	6.03*
Male	6.00*	2.81	6.01*



Gipps	Mean	Standard Deviation	Median
Female	6.35	2.66	6.11
Male	6.27	2.77*	6.22



IDM	Mean	Standard Deviation	Median
Female	6.37	2.64*	6.11
Male	6.33	2.89	6.44

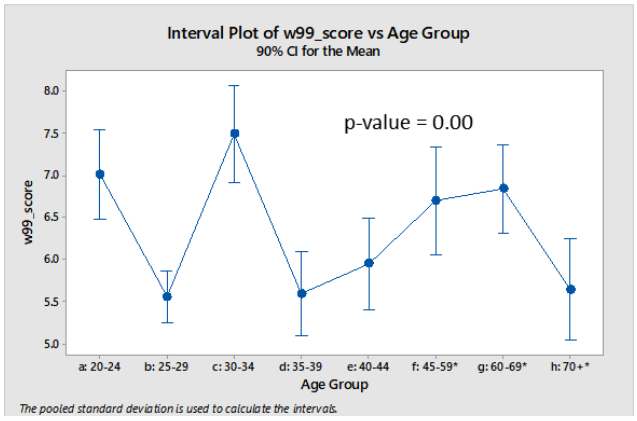
* Notates lowest value across a subcategory

Figure 6.1 Model Performance Across the Gender Attribute

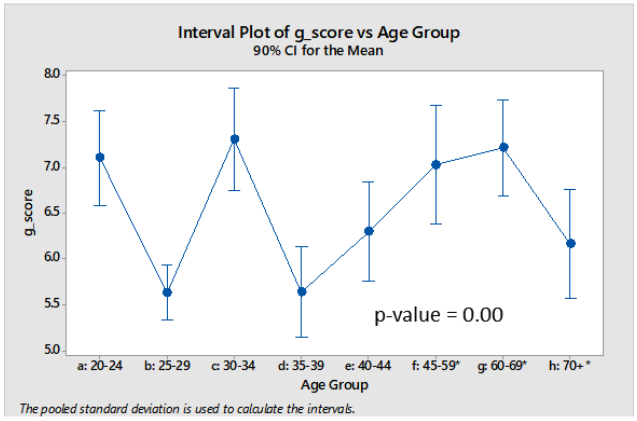
Figure 6.2 portrays the W99, Gipps, and IDM calibration scores across subcategories of age. For all three models, the 25–29 and 35–39 age groups exhibited the behavior that was the easiest to predict (i.e., lowest calibration score); the behavior of the 30–34 age group was the most challenging for the models to replicate (i.e., highest calibration score). The average and median scores varied across subcategories at a statistically significant level for all three car-following models. According to median score, the W99 best captured the driving behavior of six of the eight subcategories of age; the Gipps model best captured the driving behavior of 35–39-year-olds, while the IDM best captured the behavior of 30–34-year-old drivers.

Figure 6.3 shows the W99, Gipps, and IDM calibration scores across subcategories of race. Across all three models, the behavior of Caucasian drivers was better replicated than the behavior of drivers that did not identify as Caucasian at a statistically significant level. The W99 car-following model better captured the behavior of both subcategories of race compared to the Gipps and IDM car-following models. The W99 model also produced the highest variability in model calibration score; the Gipps model achieved the smallest within subcategory variation for both subcategories of race.

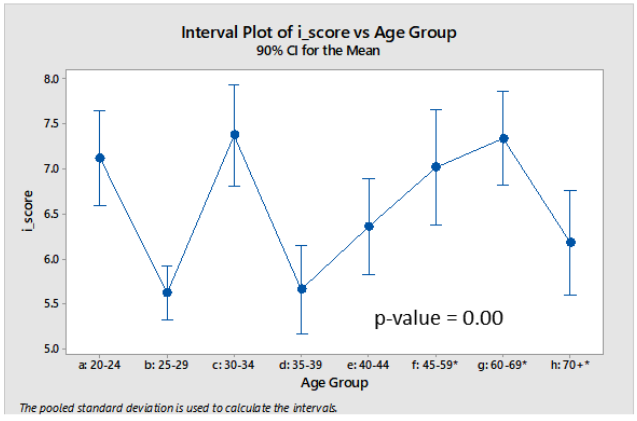
Figure 6.4 illustrates how the W99, Gipps, and IDM calibration scores vary across educational attainment. All three models were able to capture the behavior of the no college degree subcategory better than the college degree and graduate degree subcategories; the college degree subcategory's driving behavior received the highest calibration score. The score varied across educational attainment subcategories at a statistically significant level. The W99 model attained the lowest median score for the no college degree and graduate degree subcategories, while the Gipps model achieved the lowest median score for the college degree subcategory. The Gipps model obtained the lowest intra-subcategory score variation for all three car-following models.



W99	Mean	Standard Deviation	Median
20-24	7.01*	3.21	6.39*
25-29	5.56*	2.37	5.01*
30-34	7.49	2.38	6.64
35-39	5.59*	2.71*	5.50
40-44	5.95*	2.00	6.11*
45-59	6.70*	2.37	6.79*
60-69	6.84*	3.26	6.82*
70+	5.64*	3.50*	5.24*



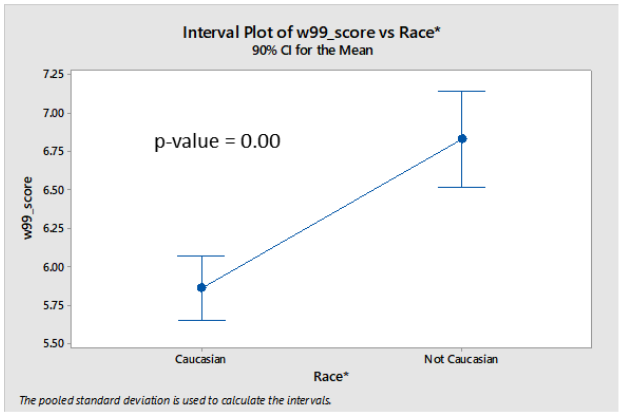
Gipps	Mean	Standard Deviation	Median
20-24	7.10	3.12	6.42
25-29	5.63	2.29	5.26
30-34	7.31*	2.21*	6.89
35-39	5.64	2.83	5.33*
40-44	6.30	1.91*	6.33
45-59	7.03	2.14*	7.26
60-69	7.21	3.16*	7.00
70+	6.16	3.51	5.86



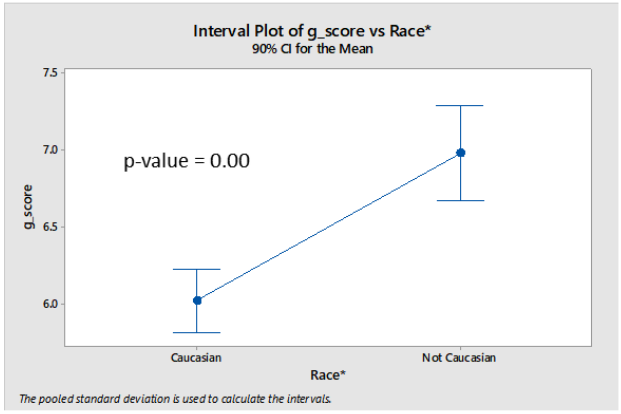
IDM	Mean	Standard Deviation	Median
20-24	7.12	2.99*	6.72
25-29	5.62	2.17*	5.43
30-34	7.37	2.40	6.61*
35-39	5.66	2.71*	5.57
40-44	6.36	2.07	6.57
45-59	7.02	2.64	6.83
60-69	7.34	3.22	7.32
70+	6.19	3.72	6.56

* Notates lowest value across a subcategory

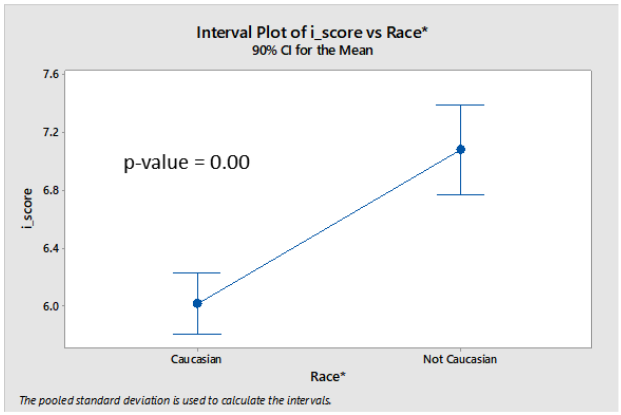
Figure 6.2 Model Performance Across the Age Attribute



W99	Mean	Standard Deviation	Median
Caucasian	5.86*	2.83	5.67*
Not Caucasian	6.83*	2.51	6.50*



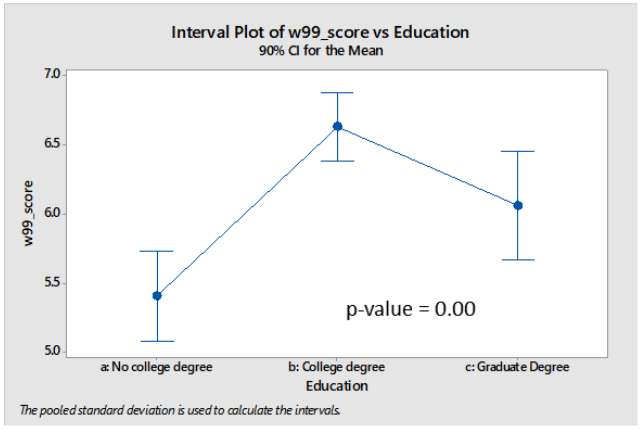
Gipps	Mean	Standard Deviation	Median
Caucasian	6.02	2.78*	5.75
Not Caucasian	6.98	2.36*	6.74



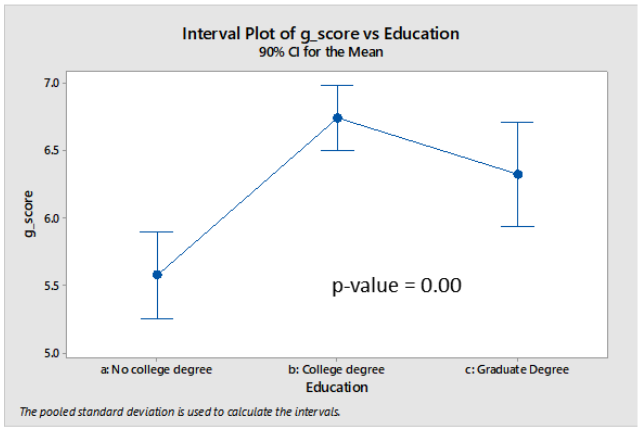
IDM	Mean	Standard Deviation	Median
Caucasian	6.02	2.79	5.78
Not Caucasian	7.08	2.42	6.93

* Notates lowest value across a subcategory

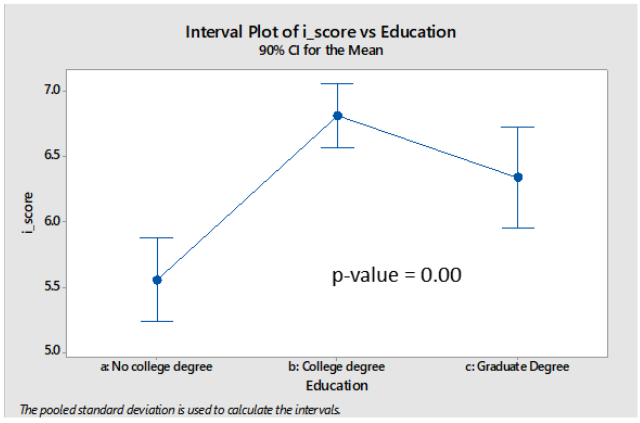
Figure 6.3 Model Performance Across the Race Attribute



W99	Mean	Standard Deviation	Median
No College Degree	5.41*	2.94	4.93*
College Degree	6.63*	2.74	6.45
Graduate Degree	6.06*	2.36	6.14*



Gipps	Mean	Standard Deviation	Median
No College Degree	5.58	2.92*	4.96
College Degree	6.74	2.68*	6.42*
Graduate Degree	6.32	2.19*	6.26



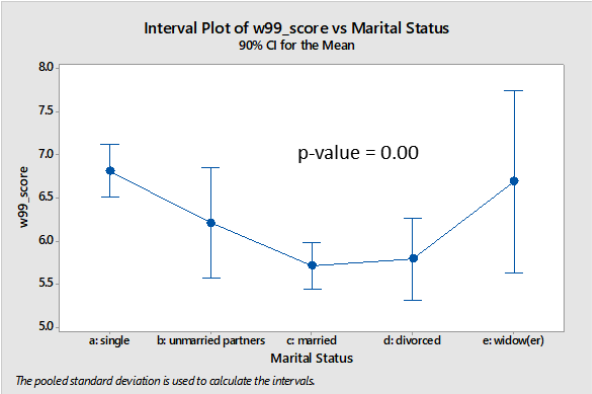
IDM	Mean	Standard Deviation	Median
No College Degree	5.55	2.97	5.02
College Degree	6.82	2.68*	6.55
Graduate Degree	6.34	2.26	6.46

* Notates lowest value across a subcategory

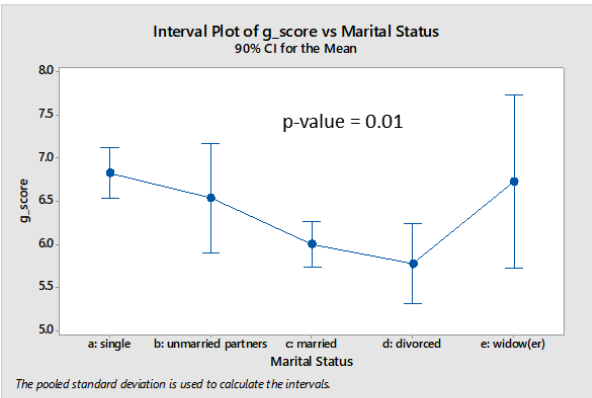
Figure 6.4 Model Performance Across the Educational Attainment Attribute

Figure 6.5 portrays how the W99, Gipps, and IDM calibration scores vary across subcategories of marital status. The W99 car-following model best captured the behavior of married drivers, while the Gipps and IDM were able to best model divorced drivers' behavior. The single and widow(er) subcategories were the hardest to replicate for the three car-following models. The variation in calibration score across the subcategories of marital status was statistically significant. The W99 model achieved the lowest median score for the unmarried partners, married, and divorced subcategories of drivers. The Gipps model best modeled single drivers, while the IDM performed best for the widow(er) subcategory of drivers. The W99 model achieved the lowest within subcategory variation for the widow(er) subcategory. The Gipps model achieved the lowest within subcategory variation for the unmarried partners and married subcategories. The IDM obtained the lowest score standard deviation for the single drivers subcategory. The W99 and IDM were able to achieve equally small score standard deviations for the divorced subcategory.

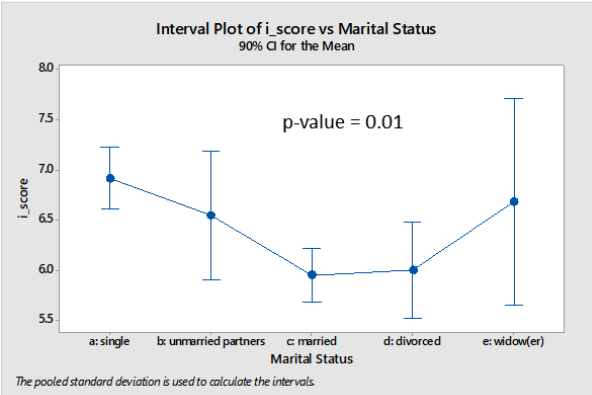
Figure 6.6 documents how the W99, IDM, and Gipps calibration scores vary across annual income. The variation in scores was statistically significant for all three models. The $50–69k income subcategory achieved the lowest calibration score for each of the three models, while the largest income subcategory reported the highest calibration score. The W99 model achieved the smallest median calibration score for half of the subcategories: $40–49k, $70–99k, and $100–149k. The Gipps model achieved the smallest median score for the under $39k and $50–69k income subcategories. Finally, the IDM model was best able to capture the driving behavior of the largest income subcategory. The Gipps model achieved the lowest within subcategory variation for all of the subcategories except one; the IDM produced the smallest within subcategory variation for the smallest income subcategory.



W99	Mean	Standard Deviation	Median
Single	6.85	2.78	6.48
Unmarried Partners	6.21*	1.60	6.27*
Married	5.71*	2.74	5.47*
Divorced	5.79	2.74*	5.50*
Widow(er)	6.69	3.98*	6.88



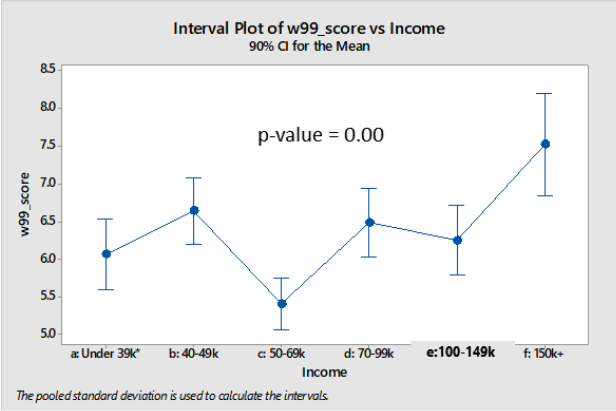
Gipps	Mean	Standard Deviation	Median
Single	6.83*	2.60	6.45*
Unmarried Partners	6.54	1.48*	6.42
Married	6.00	2.68*	5.75
Divorced	5.78*	2.97	5.63
Widow(er)	6.73	4.03	6.89



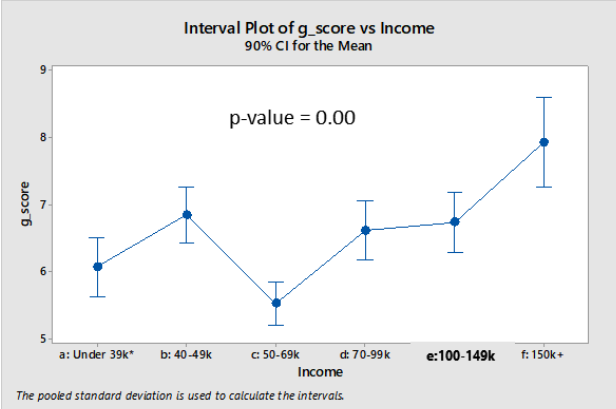
IDM	Mean	Standard Deviation	Median
Single	6.92	2.55*	6.67
Unmarried Partners	6.55	1.70	6.50
Married	5.95	2.84	5.74
Divorced	6.00	2.74*	5.80
Widow(er)	6.68*	4.26	6.66*

* Notates lowest value across a subcategory

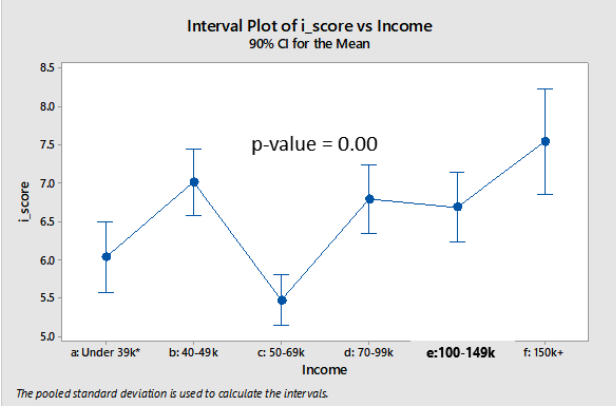
Figure 6.5 Model Performance Across the Marital Status Attribute



W99	Mean	Standard Deviation	Median
Under 39k	6.06	2.56	6.12
40-49k	6.64*	2.24	6.46*
50-69k	5.40*	2.67	4.96
70-99k	6.52	3.32	6.26*
100-149k	6.25*	2.50	6.30*
150k+	7.55	3.27	7.06



Gipps	Mean	Standard Deviation	Median
Under 39k	6.07	2.37	5.86*
40-49k	6.76	2.09*	6.70
50-69k	5.52	2.64*	4.88*
70-99k	6.48*	2.94*	6.41
100-149k	6.73	2.33*	6.59
150k+	7.53*	2.26*	7.22



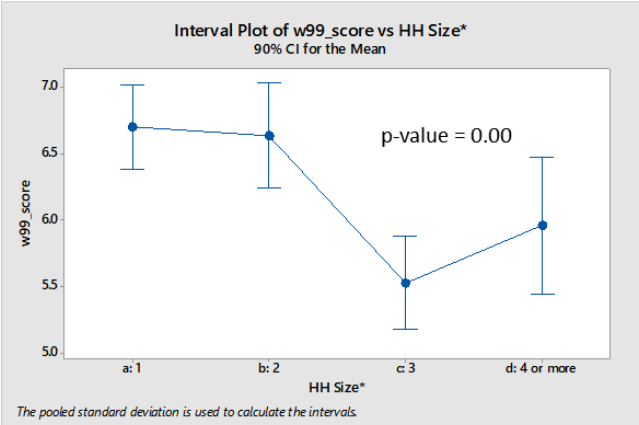
IDM	Mean	Standard Deviation	Median
Under 39k	6.03*	2.25*	6.01
40-49k	7.01	2.30	6.98
50-69k	5.47	2.68	5.00
70-99k	6.79	3.22	6.64
100-149k	6.68	2.60	6.67
150k+	7.54	2.88	6.99*

* Notates lowest value across a subcategory

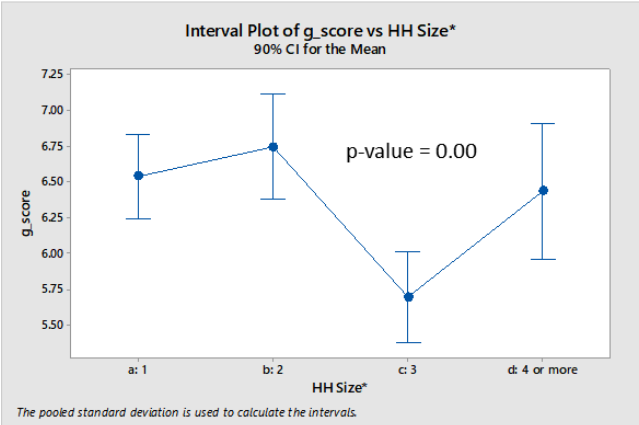
Figure 6.6 Model Performance Across the Annual Income Attribute

Figure 6.7 shows how the W99, IDM, and Gipps calibration scores vary across reported household size. The variation in scores was statistically significant for all three models. The three household members subcategory recorded the lowest calibration score for each of the three models. The W99 model best replicated the driving behavior of each subcategory of data. Moreover, the W99 model achieved the lowest within subcategory variation for all subcategories except the single household member subcategory.

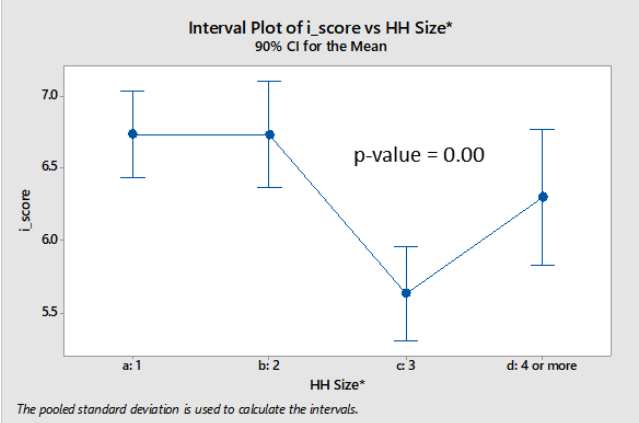
Figure 6.8 illustrates how the W99, IDM, and Gipps calibration scores vary across estimated mileage driven over the last calendar year. The variation in scores was statistically significant across the subcategories for all three models. Reported driver mileage and calibration scores were inversely correlated, with the lowest category of mileage achieving the highest calibration score and the highest category of mileage achieving the lowest calibration score. The W99 model achieved the smallest median calibration score for the 0–5k, 10–12k, and 20–23k subcategories. The Gipps model best matched the 16–19k and 25k+ subcategories. The IDM best replicated the 6–9k and 13–15k subcategories. Overall, the driver mileage last year attribute achieved the lowest average within attribute score (i.e., average(median score) across all subcategories of driver mileage last year).



W99	Mean	Standard Deviation	Median
1	6.45*	2.78	6.34*
2	6.49*	2.41*	6.46*
3	5.55*	2.44*	5.27*
4+	6.05*	2.63*	6.07*



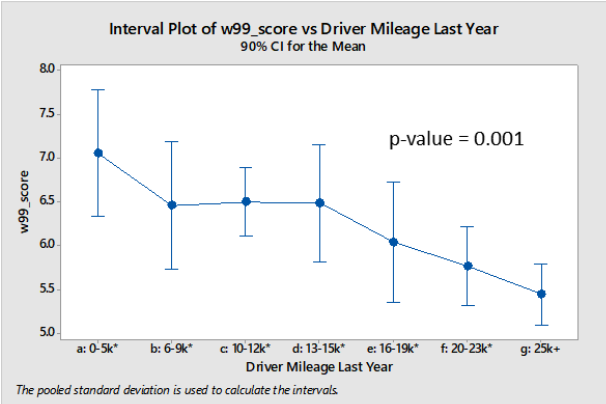
Gipps	Mean	Standard Deviation	Median
1	6.54	2.87	6.36
2	6.74	2.57	6.64
3	5.69	2.50	5.47
4+	6.43	2.71	6.33



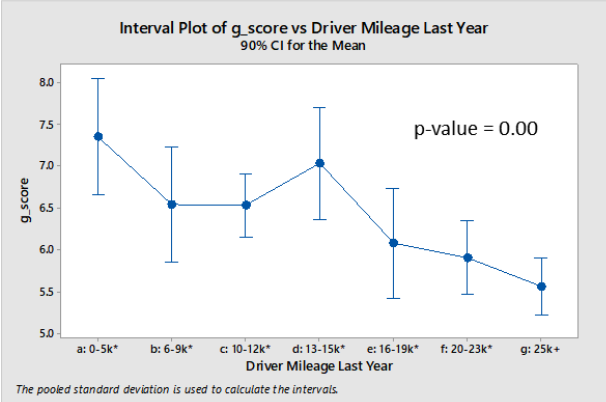
IDM	Mean	Standard Deviation	Median
1	6.74	2.76*	6.56
2	6.73	2.70	6.66
3	5.63	2.49	5.46
4+	6.30	3.00	6.46

* Notates lowest value across a subcategory

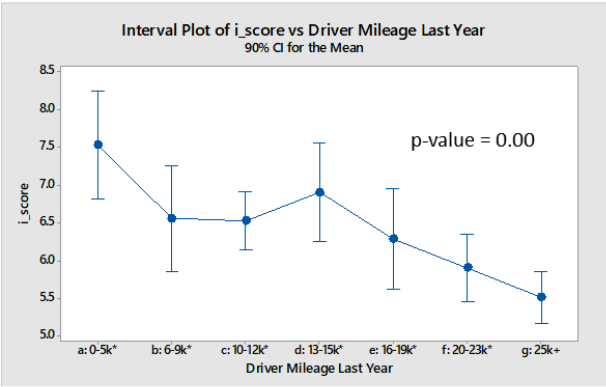
Figure 6.7 Model Performance Across the Household Size Attribute



W99	Mean	Standard Deviation	Median
0-5k	7.06*	2.44	6.71*
6-9k	6.46*	3.68*	6.41
10-12k	6.50*	2.62	6.20*
13-15k	6.49*	3.99	7.05
16-19k	6.04*	2.80	5.30
20-23k	5.77*	2.89*	5.66*
25k+	5.44*	2.22	5.38



Gipps	Mean	Standard Deviation	Median
0-5k	7.36	2.47	6.91
6-9k	6.54	3.72	5.80
10-12k	6.53	2.46	6.26
13-15k	7.04	3.62*	7.25
16-19k	6.08	2.54	5.08*
20-23k	5.90	2.93	5.83
25k+	5.56	2.16	5.34*



IDM	Mean	Standard Deviation	Median
0-5k	7.53	2.20*	7.32
6-9k	6.56	3.89	5.32*
10-12k	6.53	2.32*	6.40
13-15k	6.90	4.18	6.94*
16-19k	6.28	2.51*	6.11
20-23k	5.90	2.90	6.07
25k+	5.51	2.15*	5.45

* Notates lowest value across a subcategory

Figure 6.8 Model Performance Across the Reported Annual Mileage Driven Last Year Attribute

6.4. CHAPTER 6 CONCLUSIONS

Chapter 6 explored the variability of model performance across different subcategories of drivers segmented by their driver specific attributes. Towards answering research question number 3 posted by this dissertation, Chapter 6 indicates that different car-following models have varying degrees of success in capturing the driving behavior of different drivers. This is evidence that different car-following models should be used to approximate the behavior of different drivers. Furthermore, Chapter 6 suggests that driver attributes are one element that could be used successfully to decide which model should be used to characterize the behavior of a driver.

The key points of the chapter are summarized as follows:

- The variation in score was statistically significant across all driver attribute subcategories except those that comprise the gender attribute. This provides support that different subcategories of driver attributes should be modeled differently.
- No one model best described all subcategories of data. This further supports previously documented observations that in order to adequately capture the heterogeneity in naturalistic driving data, different sets of optimal calibration parameters and different car-following models may be required (Ossen, 2008; Ossen & Hoogendoorn, 2005, 2011).
- The driver mileage last year attribute attained the lowest average score across subcategory medians (i.e., the average of the median scores for the 0–5k, 6–9k, 10–12k, 13–15k, 16–19k, 20–23k, and 25k+ subcategories).
- The educational attainment attribute obtained the lowest average score across subcategory means (i.e., the average of the mean scores for the no college degree, college degree, and graduate degree subcategories).

- The annual income driver attribute obtained the lowest average standard deviation across the subcategory scores (i.e., the average of the standard deviations of the under $39k, $40–49k, $50–69k, $70–99k, $100–149k, and $150k+ income subcategories).

As a recap, Chapter 5 established that the behavior of different subgroups of drivers, defined by subcategories of driver attributes, are statistically significant; moreover, the application of clustering algorithms has indicated that there are homogeneous groups of drivers in the naturalistic data, which can be defined through driver attributes. Practically speaking, this means that behavioral heterogeneity can be reduced by controlling for driver specific attributes. Chapter 6 provided evidence that no one model best describes all of the drivers in the sample of SHRP2 NDS; that is, some models describe subgroups of drivers better than others. The next chapter of this dissertation seeks to develop a method to obtain a set of representative car-following model calibration coefficients for collections of diverse trajectories (i.e., male drivers, young drivers, married drivers).

Chapter 7: Methods to Obtain Representative Car-Following Model Parameters from Trajectory-Level Data for Use in Microsimulation (Task 3)[2]

In Chapters 5 and 6, support was provided that suggests that drivers with different driver attributes (e.g., young female drivers, middle-aged female drivers, older male drivers) may need to be modeled differently (i.e., different car-following model calibration parameter estimates; different car-following models) to more adequately capture the inter-driver heterogeneity evident in naturalistic driving data. However, to the author's best knowledge, minimal research has been conducted to identify methods to aggregate diverse trajectories from groups of drivers (i.e., the homogeneous driver groups discussed in Chapter 5) into one representative set of model parameters. This chapter explores eight viable methods for obtaining representative sets of calibration parameter coefficients for a group of drivers and recommends a best method for the Wiedemann 99 (W99), Gipps, Intelligent Driver Model (IDM), and Newell car-following models. A paper describing this effort was submitted for presentation at the 2019 Annual Meeting of the Transportation Research Board and for publication in the Transportation Research Record. The paper was accepted for presentation and is currently in a second round of reviews for publication. A citation for this paper is as follows:

James, R. M., Hammit, B. E., and Boyles, S. D. (2018). Methods to Obtain Representative Car-Following Model Parameters from Trajectory-Level Data for Use in Microsimulation. *In Press: Transportation Research Record*.

[2] Chapter 7 will be featured in an upcoming publication of the Transportation Research Record. The full citation is as follows. James, R. M., Hammit, B. E., and Boyles, S. D. (2018). Methods to Obtain Representative Car-Following Model Parameters from Trajectory-Level Data for Use in Microsimulation. *In Press: Transportation Research Record*. R. M. James and B. E. Hammit contributed nearly equally to the development of the research idea, the modeling and interpretation of results, and the development of the paper. S. D. Boyles developed the mathematical notation for the publication of the paper.

7.1. MOTIVATION

Transportation agencies are seeking reliable methods to reduce the burden of resource constraints through improved decision-making. Microsimulation models are one such tool available to agencies. These models facilitate detailed scenario analyses, enabling agencies to obtain robust estimates of performance metrics for possible intervention strategies; this significantly improves their ability to make decisions by comparing estimated costs with a more realistic estimate of benefits. However, these detailed models are highly dependent both on the realism of the underlying sub-models (e.g., car-following, lane-changing) and the accuracy of the input data used for model calibration. Of particular importance is the validity of car-following models, the sub-model within microsimulation that controls longitudinal driving behavior (i.e., acceleration).

Emerging sources of trajectory-level data are providing researchers with an unprecedented opportunity to improve car-following models. To date, this has primarily been studied through the process of calibration, where trajectory-level data are used to obtain optimal parameter sets for common car-following models. There is now consensus in the literature regarding best practices to obtain a calibrated parameter set for a single collection of data (e.g., a single driver's trajectory; a single trip) (Ciuffo et al., 2012; Hammit et al., 2018a; Montanino et al., 2012; Punzo & Montanino, 2016; Treiber & Kesting, 2013b). However, the application of trajectory-level data is challenging common practices for using car-following models in microsimulation, particularly the assumption of behavioral homogeneity within a designated driver population. Through the calibration of car-following models, several researchers have noted the existence of differences in driving behavior, both between different drivers and within a single driver, evident in trajectory-level data (Brockfeld et al., 2004; Ossen & Hoogendoorn, 2005, 2011; Ossen et al., 2006; Punzo & Simonelli, 2005; Sangster et al., 2013; Soria et al., 2014). Moreover, it

has been shown that this behavioral heterogeneity impacts the propagation of shockwaves (Ossen & Hoogendoorn, 2008) and ultimately has an impact on the predicted capacity of a facility (Ossen & Hoogendoorn, 2011).

Therefore, methods to practically account for behavioral heterogeneity deserve additional attention. For the application of trajectory-level data in the model calibration process to become feasible in practice, methods to sample the observed driving behaviors to obtain representative parameter sets must be reliably demonstrated. The adoption of these methods to characterize a population of drivers (e.g., men, younger drivers) or driving conditions (e.g., snowy weather, work zones, narrow lanes) have been impeded by ambiguous validation results.

Validation practices in transportation often involve the use of holdout data: can the calibrated parameter sets obtained on a trajectory from dataset A be used to accurately predict a trajectory from dataset B, which was not used in the calibration process? (Brockfeld & Wagner, 2006; Punzo & Simonelli, 2005; Zheng et al., 2012). However, this procedure has consistently resulted in high validation errors regardless of the car-following model used. As a result, many studies associate the high errors as evidence of the model overfitting to the behavior observed in the calibration data (Brockfeld et al., 2004; Punzo & Simonelli, 2005; Zheng et al., 2012).

Although the hold-out validation procedure is ubiquitously accepted, the data science literature provides additional insights into alternative procedures that result in robust error estimates while reducing the risk of overfitting (Witten, Frank, Hall, & Pal, 2017a). In particular, the n-fold cross-validation procedure, or the repeated holdout method of error estimation, allows modelers to use all available data for training, while producing a reliable estimate of the error; this reduces the likelihood of using an unrepresentative dataset for training or validation.

This paper applies a ten-fold cross-validation framework to a 100-trip sample from the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) to test eight methods for obtaining representative car-following model parameters that describe a collection of trips. These eight methods build on prior work (Hammit et al., 2018a) focusing on the ultimate goal of practical implementation. The methods are divided into two groups: methods that preserve possible correlations between parameter coefficients (Kim & Mahmassani, 2011; Monteil, Billot, Sau, Buisson, & El Faouzi, 2014) and methods that obtain parameter estimates independent of the other parameter coefficients. Intuitively, the methods adopt the following strategies: (i) identifying average behavior, (ii) identifying the most frequently occurring behavior, and (iii) using random sampling techniques.

The remainder of this chapter is organized as follows. Section 7.2 discusses the data sampling procedure used to obtain a 100-trip sample for the SHRP2 NDS dataset and the trajectory-level calibration procedure used to obtain the trip-specific sets of calibration parameter coefficient estimates for each car-following model. Section 7.3 discusses the eight viable methods for obtaining a representative set of calibration coefficients and the Calibration-Validation framework used to evaluate the sets of calibration coefficients for each method. The cross-validation results are discussed in Section 7.4. Finally, concluding remarks for this chapter are provided in Section 7.5.

7.2. DATA ACQUISITION AND SAMPLING

This section uses a sample of the SHRP2 NDS dataset queried through the Wyoming Department of Transportation (WYDOT) Implementation Assistance Program (IAP); for additional details on the SHRP2 NDS and the WYDOT IAP, see Section 3.1.1

and Section 3.1.3, respectively. The goal of this paper is to evaluate methodologies to obtain a representative car-following model parameter set to describe a population of drivers or specific driving condition. Therefore, the WYDOT sample was further reduced to remove unnecessary sources of heterogeneity that could confound validation results. To this end, the road type and weather condition were restricted to freeways and clear weather conditions, as past studies have shown these factors are sources of intra-driver heterogeneity. Moreover, to ensure sufficient data for calibration, sampled trips are at least 1000 seconds in length (i.e., at least 10,000 inter-vehicle spacing data observations available for each trip).

A distinction between instrumented research vehicle (IRV), such as those collected via SHRP2, and aerially collected trajectories is the spatiotemporal scope. Aerially collected trajectories tend to be limited in scope (e.g., 1000-m in the I–80 NGSIM dataset), while IRV trajectories cover the entire duration of a driver's trip. Some of the trajectories in the WYDOT sample contain over an hour of continuous data, only some of which is defined as car-following. A radar-vision algorithm to identify the presence of a leading vehicle through CAN-Bus and radar data was developed in a previous effort (Hammit et al., 2018b) and is described in Section 3.2.2. This algorithm was applied to all WYDOT trips to identify the car-following segments in a trip. It should be noted that only constrained driving states were used for calibration. However, it was confirmed that the average maximum following distance for the constrained driving states is about 90 m; this distant "leader" should allow for calibration of parameters that represent free driving states, such as desired velocity. Moreover, a data completeness framework was developed based on the definitions of driving regimes in the W99 car-following model (i.e., free driving, approaching, following, crash zone) to confirm that sufficient data in each of the four regimes was evident in the queried dataset.

Ultimately, a 100-trip sample was randomly obtained from the set of data that met the road type, weather, and trip length thresholds. Summary statistics for the 100-trip sample are shown in Table 7.1. Each of the four models—W99, Gipps, IDM, and Newell—were calibrated independently to match the trajectories of each of the 100 trips. This produced 400 sets of calibrated parameters. A set of estimated calibration parameters obtained for a specific trip will be referred to as the "trip-specific calibration parameter coefficients" for the remainder of this paper.

Table 7.1 Descriptive Statistics for 100-Trip Sample

Descriptive Statistics	**Minimum**	**Average**	**Maximum**	**Standard Deviation**
Trip Duration [min]	16.8	22.7	33.2	2.80
Number of Driving States (i.e., car-following or non-car-following)	10.0	25.6	65.0	12.7
Time Spent in Constrained Driving State (i.e., car-following) [min]	8.54	11.8	19.5	4.40
Velocity in Constrained Driving State (i.e., car-following) [m/s]	17.9	27.7	34.9	3.4
Time Gap in Constrained Driving State (i.e., car-following) [s]	0.70	1.96	4.95	0.66
Distance Traveled During Constrained Driving State (i.e., car-following) [m]	80.0	371	1585	218
Following Distance in Constrained Driving State (i.e., car-following) [m]	20.9	53.0	150	17.1

The calibration procedure was derived from a survey of best practices in calibrating car-following models using trajectory-level data, which is documented in Section 3.2.4. It was identified that optimization problems designed to minimize the root mean square error (RMSE) between the predicted and observed following distance profiles between a leader-follower pair of vehicles (dX) achieves the most reliable calibration results. Specifically, let m index the car-following models (i.e., trajectory translation, safety distance, social force, and psychophysical). Each model involves parameters θ_m which must be calibrated; let Θ_m be the set of all such parameters for model m and $\overline{\Theta_m}$ represent the set of feasible parameters for this model. Let i index the trips in the dataset and t index the time stamps for the observations. The dataset contains $dX_i(t)$ values, denoting the following distance to the lead vehicle for each trip and time index. Given the trajectory of the lead vehicle, and given parameters Θ_m, each model m produces an estimate $\widehat{dX}_{i,m}(t|\Theta_m)$ of this distance.

Trip-specific calibration parameter estimates were obtained by minimizing RMSE of these following distances, producing trip-specific calibration parameters $\Theta^*_{i,m}$ for each model, that is,

$$\Theta^*_{i,m} \in \arg\min_{\Theta \in \overline{\Theta_m}} \sqrt{\sum_t (dX_i(t) - \widehat{dX}_{i,m}(t|\Theta)} \quad 5.1$$

From the calibration process, 100 "trip-specific calibration parameter coefficients" were obtained for each model.

7.3. CALIBRATION-VALIDATION FRAMEWORK

This section provides an overview of the eight calibration methods and discusses the 10-fold cross-validation procedure used to evaluate the performance of each method.

The goal of these methods is to obtain a set of parameters Θ_m^* for a particular model across *all* trips, in contrast to the $\Theta_{i,m}^*$ values, which are specific to a particular trip (Section 3.2.4).

7.3.1. Proposed Calibration Methods

This study evaluates eight different methods for obtaining a representative set of car-following parameters to describe a population of drivers or a specific driving condition. The differences between each of the eight methods are visible in Figure 7.1.

The methods can be categorized in several ways. Method 1 is truly a "calibration" method, as it does not apply the trip-specific calibrated parameter sets; rather, it ingests the raw car-following data from all trips and solves a nonlinear optimization problem to obtain a single set of calibration parameters. Methods 2–8 are more accurately described as sampling methods, as they sample from the trip-specific parameter coefficients. Although Method 2 makes use of the previously obtained trip-specific parameter coefficients, it also requires the original trajectory-level driving data, as discussed next. Methods 3–8 only require the sets of optimal trip-specific parameter coefficients $\Theta_{i,m}^*$ that were previously identified, which makes these methods easier to apply, especially for large datasets.

The methods also differ in whether they maintain possible correlations between parameters (Kim & Mahmassani, 2011; Monteil et al., 2014). Many car-following model parameters have physical interpretations, such as desired acceleration or reaction time (Gipps, 1981). However, when calibrated alongside other model parameters, the values may not reflect this intended meaning due to influence by other parameters. For example, interaction between the computed reaction time and desired acceleration parameters of the Gipps model suggest that observed trajectories could be described either by high acceleration rates with low reaction times, or low acceleration rates with high reaction

times. Methods 1–4 maintain the presumed relationships between calibrated parameter coefficients, while Methods 5–8 relax this assumption. Finally, all methods have qualitative interpretations corresponding to alternative calibration goals: Methods 1, 2, 5, and 6 represent different ways to obtain "average" driving behavior; Methods 3 and 7 are intended to capture the most frequently occurring driving behavior; and Methods 4 and 8 attempt to capture inherent variability in driving behavior through random sampling techniques.

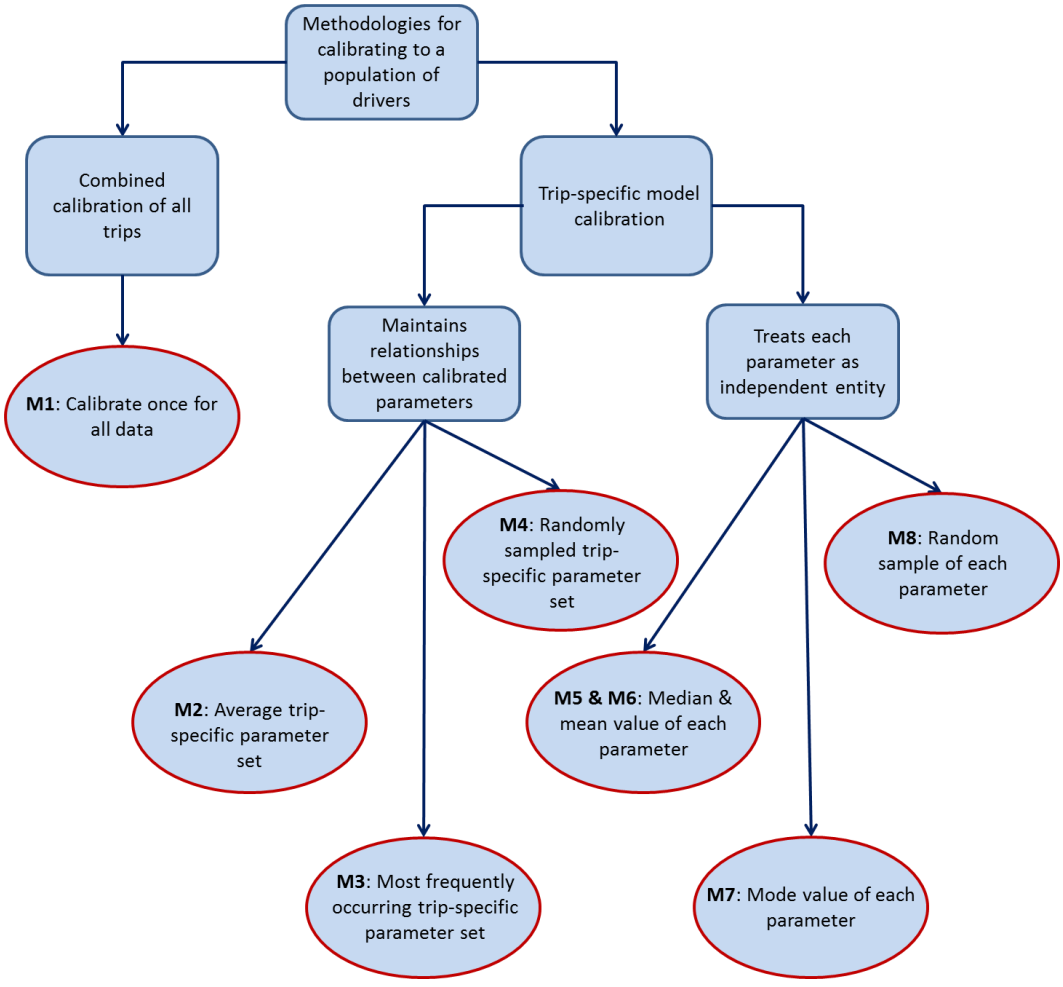


Figure 7.1 Eight Proposed Calibration Methods used to obtain a Representative Set of Model Parameters (James, Hammit, & Boyles, 2019)

Methods Requiring Original Trajectory-Level Data

Methods 1 and 2 are the most computationally expensive, because they require the original trajectory-level data to identify representative calibration parameters. Method 1 should be thought of as a calibration procedure, as it is the only method that requires the use of a genetic algorithm to generate an optimized set of parameter coefficients, Θ_m^*. Method 2 is a systematic sampling procedure, using the trip-specific calibration coefficients $\Theta_{i,m}^*$ obtained from Equation 5.1. It is less computationally expensive than Method 1 because it leverages the trip-specific calibration coefficients, which are assumed to have been previously computed. However, Method 2 is more computationally expensive than Methods 3 through 8 because it requires the use of the trajectory-level data to evaluate how well an existing trip-specific set of estimated parameter coefficients predicts the behavior observed in the calibration dataset. Method 1 and Method 2 are described in more detail in the following subsections.

Method 1

Method 1 is the most computationally intensive method, as it uses all of the available segments of car-following data in a single calibration process to obtain one optimal parameter set. The calibration procedure used for Method 1 was described in Section 7.2, the difference being the amount of data used in the calibration process (i.e., all the available data vs. the data for a single trip). The Method 1 representative parameter set is the set that best minimizes the weighted average RMSE(dX) across all of the car-following states in the calibration data simultaneously:

$$\Theta_m^* \in \arg\min_{\Theta \in \Theta_m} \sqrt{\sum_i \sum_t (dX_i(t) - \widehat{dX}_{i,m}(t|\Theta)} \qquad 5.2$$

A genetic algorithm was used to solve this optimization problem.

Method 2

Method 2 identifies the "average" behavior observed in the calibration data, while maintaining possible correlations between parameters. This is the most computationally expensive method that makes use of the previously obtained sets $\Theta_{i,m}^{*}$ (Methods 2–8). However, it is less burdensome than Method 1 because it does not optimize over the space of *all* feasible parameter coefficients; instead, it only searches the previously obtained trip-specific parameter sets $\Theta_{i,m}^{*}$, identifying the trip-specific parameter set best matching all trips in the calibration dataset:

$$\Theta_{m}^{*} \in \arg\min_{i} \sqrt{\sum_{i}\sum_{t}(dX_i(t) - \widehat{dX}_{i,m}(t|\Theta_{i,m}^{*})} \quad 5.3$$

The optimal solution is found by enumerating over the trip-specific parameter sets. Specifically, the following logic is applied: (i) identify a set of trip-specific calibration parameters in the calibration data; (ii) calculate the error (i.e., RMSE(dX)) in using that trip-specific parameter set to describe all the car-following states in the calibration data; (iii) repeat steps 1 and 2 using every set of trip-specific calibration parameters in the calibration dataset; and (iv) identify the set of estimated trip-specific calibration parameters that produce the minimum RMSE(dX) across all of the calibration data. This trip-specific parameter set was selected as the Method 2 representative parameter set.

Methods Not Requiring Original Trajectory-Level Data

Method 3–8 are less computationally expensive, as they only involve the trip-specific parameter sets $\Theta_{i,m}^{*}$ and not the original trajectory data. Method 3 through Method 8 assume that the trip-specific calibration coefficients, as described in Section 7.2, have already been obtained and are available for sampling.

Method 3

Method 3 captures the most frequently observed driving behavior, maintaining the relationships and possible correlations between estimated parameter coefficients:

$$\Theta_m^* = \text{mode}\{\Theta_{i,m}^*\} \quad 5.4$$

In the experiments that follow, ties are broken randomly. In the case where the $\Theta_{i,m}^*$ are all distinct, the method fails to produce a value. Such occurrences are relatively rare and are noted in the numerical results.

Method 4

Method 4 produces a set of parameters by sampling from the complete set of trip-specific calibration parameter coefficients:

$$\Theta_m^* = \text{mode}\{\Theta_{i,m}^*\} \quad 5.5$$

This is intended to capture the common practice of randomly sampling parameter coefficients from a desired distribution, while maintaining the potential relationships between parameter coefficients. The experimental results evaluate this method by averaging over ten samples from the set of $\Theta_{i,m}^*$ values.

Methods 5–8

The remaining methods identify parameter coefficients θ_m independently of the other model parameters, rather than estimating the entire set Θ_m at once. In what follows, the notation $\theta_{i,m}^*$ refers to the specific value of the parameter θ_m from the trip-specific calibration parameter set $\Theta_{i,m}^*$. The methods differ according to how the parameter is chosen: Method 5 chooses the median (Equation 5.6), Method 6 the mean (Equation 5.7), Method 7 the mode (Equation 5.8), and Method 8 a random sample (Equation 5.9), like what was defined in Method 4.

$$\theta_m^* = \text{median}\{\theta_{i,m}^*\} \quad 5.6$$

$$\theta_m^* = \text{mean}\{\theta_{i,m}^*\} \quad 5.7$$

$$\theta_m^* = \text{mode}\{\theta_{i,m}^*\} \quad 5.8$$

$$\theta_m^* = \text{sample}\{\theta_{i,m}^*\} \quad 5.9$$

7.3.2. Ten-Fold Cross-Validation Procedure

This study uses *n*-fold cross-validation to evaluate the methods. This procedure iteratively changes which portion of the data is used in the validation process (Witten et al., 2017a). This method is able to ascertain a fair assessment of model error by always ensuring the data used for model validation are not used in the calibration process. This method also obviates the decision of how much data to use in training and testing, by iteratively using each fold of data. In data science, cross-validation is often applied as follows. Using nine folds of the data (e.g., F2–10), a model is learned; this learned model is evaluated using the first fold of data, which was withheld from the learning process and a validation score is recorded. The next iteration holds out a new fold of data (F2) while the remaining nine folds (F1,3–10) are used to learn a model. This process is repeated using each fold, obtaining 10 different validation scores using the 10 different sets of data iteratively withheld from the learning process. To assess the overall performance of the model, the validation scores are averaged. A final model is learned using all of the data and the model score is reported as the aforementioned average validation score across the 10 folds of holdout data. This method reduces the likelihood of overfitting or using an unrepresentative dataset for either calibration or validation by allowing the modeler to use all available data.

For this study, a 10-fold cross-validation procedure was used. Ten folds were selected because this number works well in multiple domains (Witten et al., 2017a). This procedure is used to assess the eight calibration methods, with the intention of identifying the method producing the smallest error estimate, or "fold score", computed as the average RMSE(dX) between the predicted behavior from the set of parameters identified in the calibration process and the actual car-following behavior in the validation fold. The "method score" is the average RMSE(dX) across all 10 folds of validation data.

This complete procedure is illustrated in Figure 7.2. The 100-trip sample is first randomized and segmented into 10 equivalent folds. Nine-folds are merged to create the calibration data (90 trips) and one-fold is isolated for validation (10 trips). The radar-vision algorithm (Hammit et al., 2018b) is applied to identify all car-following segments in the calibration and validation datasets. Then, the procedure in Hammit et al. (2018a) obtains trip-specific calibration parameters $\Theta_{i,m}^{*}$. Each of the eight methods are then applied and used to calculate the fold score. This procedure is repeated 10 times, using each of the data folds for validation. Once each fold of data has been used for validation, the method score is computed.

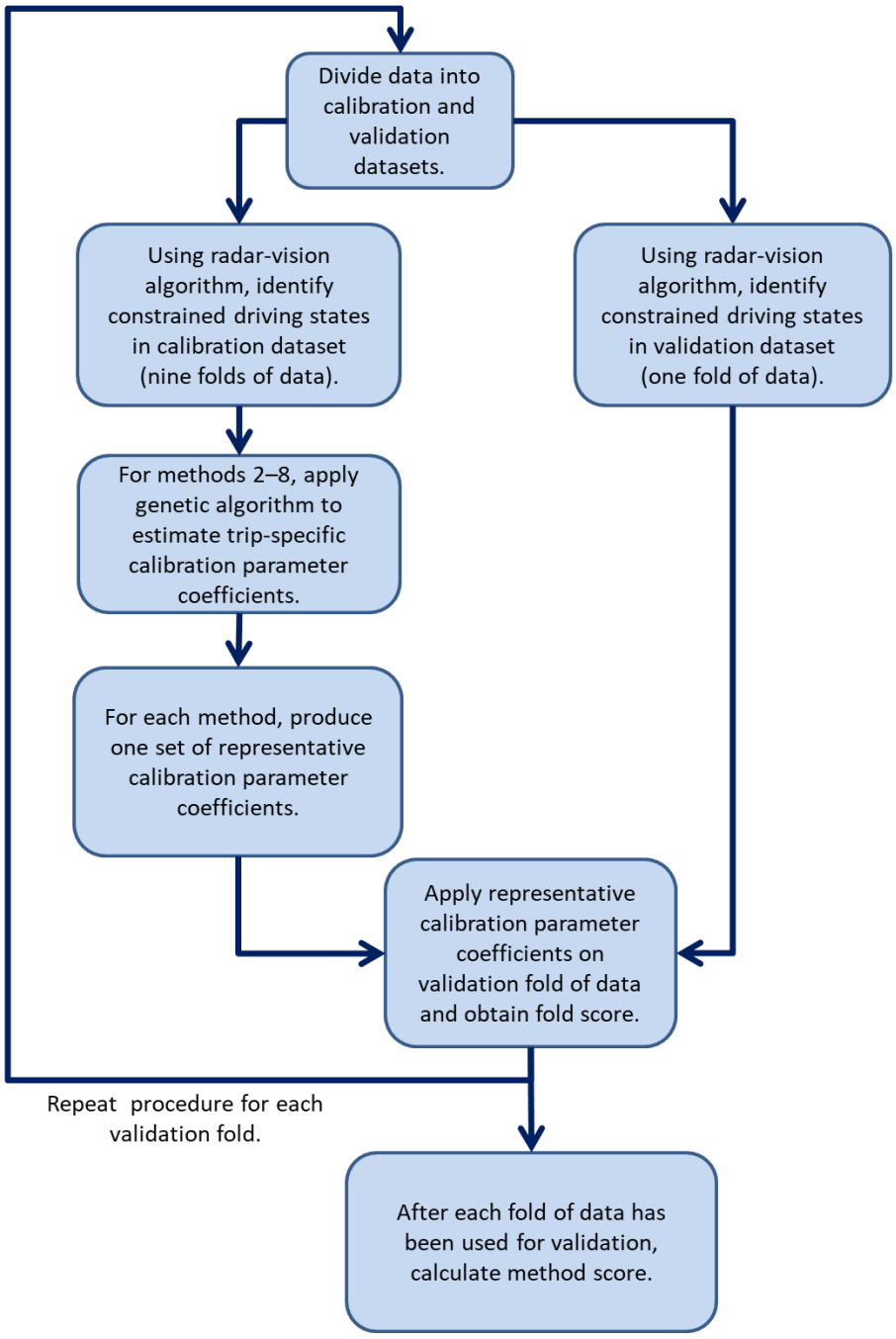


Figure 7.2 Validation Framework for Evaluating Calibration Methods (James, Hammit, & Boyles, 2019)

7.4. RESULTS

The results are organized in two sections. First, the performance of the car-following models is discussed. Second, a comparison of the calibration methods is conducted.

7.4.1. Performance of Car-Following Models

Table 7.2 provides the cross-validation results for each calibration method performed with each car-following model. When comparing models across the best performing calibration method, the results indicate that W99—the most complex car-following model with 11 calibration parameters—had the best overall performance. This finding suggests that W99, on average, can better predict the naturalistic driving behavior in the SHRP2 NDS. Conversely, Newell performed the worst, with a score 50% higher than that of W99.

The Newell model achieved the lowest standard deviation of model scores across all calibration methods, followed by IDM, W99, and Gipps. Due to the simplicity of the Newell model, it is not able to match the driving trajectories as well as the other models; this phenomenon is called underfitting. Underfitting occurs when a model cannot capture underlying trends within the input data. The Gipps model had the second-best overall score, yet the highest standard deviation; when it performed poorly, significantly larger validation scores were reported (e.g., M8). This finding is associated with the actual model theory; although Gipps and IDM contain the same number of calibration parameters, the underlying assumptions used to form the Gipps model introduce the possibility for instability caused by certain combinations of parameters within a set. When combined, these poor parameter sets produce unrealistic predictions often resulting in excessive vehicle oscillations and crashes (Wilson, 2001). This finding further supports the notion

that parameter correlation and interaction are critical factors when sampling representative parameter sets and that the influence of this correlation has a greater impact on certain car-following models.

Table 7.2 Ten-Fold Cross-Validation Results for Each Calibration Method

Calibration Methods	Validation Error Estimate Method Score			
	Gipps	IDM	Newell	W99
Number of Model Parameters	6	6	2	11
Literature (LIT) values (Gipps, 1981; Kesting et al., 2010; Liu, 2016; Punzo & Simonelli, 2005)	26.73	42.40	13.99	25.95
M1: Calibrate once for all data	9.76	8.55	13.04	9.10
M2: Average trip-specific parameter set	8.23*	8.41*	12.99*	7.70*
M3: Most frequently occurring trip-specific parameter set	---	---	13.10	---
M4: Randomly sampled trip-specific parameter set	38.07	37.51	13.75	46.60
M5: Median value of each parameter	10.41	9.43	12.99*	10.83
M6: Mean value of each parameter	9.41	8.59	13.03	11.80
M7: Mode value of each parameter	26.94	15.55	13.10	21.82
M8: Random sample of each parameter	83.64	25.14	14.35	37.79

7.4.2. Comparison of Calibration Methods

Figure 7.3 illustrates the relative ranking of each calibration method. For all four car-following models, the best performing method was Method 2, which identified the average trip-specific parameter set, respecting the potential relationships between

parameter coefficients. The only exception to this observation is seen in the Newell model, where Method 5 exhibited the same performance; further investigation showed that the average trip-specific parameter set was equivalent to the independently calculated medians. This is not surprising due to the simplicity of the Newell model. For IDM and W99, the second-best performing method was Method 1, where the calibration procedure was completed once with all calibration data. For Gipps and Newell, the second-best performing method was Method 6, which calculated the mean value for each of the parameters independently.

Each of the eight methods fit into three behavioral categories: average behavior (M1, M2, M5, M6); most frequent behavior (M3, M7); and randomly sampled behavior (M4, M8). Four of the methods preserve the relationships between the model parameters by sampling complete sets of parameter coefficients from the calibration data, while the remaining methods obtain individual parameter estimates independent of the other parameters. In most cases, the methods that preserved the unknown relationships between the model parameters (M2, M3, M4) outperformed their counterpart that treated parameters as individual entities (M5/6, M7, M8). This further adds corroborative evidence that the preservation of these unknown relationships between parameters are important to ensure sampled parameter sets produce realistic behavior that will generalize to an unobserved population.

The exceptions are observed in the random sampling methods for IDM and W99 (M4, M8). For the purpose of this case study, the reported scores for Methods 4 and 8 are the average method score calculated from 10 random samples to reduce the risk of selecting one highly unrepresentative parameter set. Regardless, Methods 4 and 8 were the worst performing methods for each model. Further investigation illustrates that these methods have the highest standard deviation, i.e., 12.2 meters (M4) and 26.4 meters (M8), compared

to the lowest standard deviation of 1.3 meters (M5). Although Method 4 preserved the correlation structure between the parameters, the completely random selection of trips resulted in the capturing of driving behavior that was not representative of the larger population; instead, methods that did not preserve the correlation structure, but did capture more representative average behavior, outperformed this method (M5, M6).

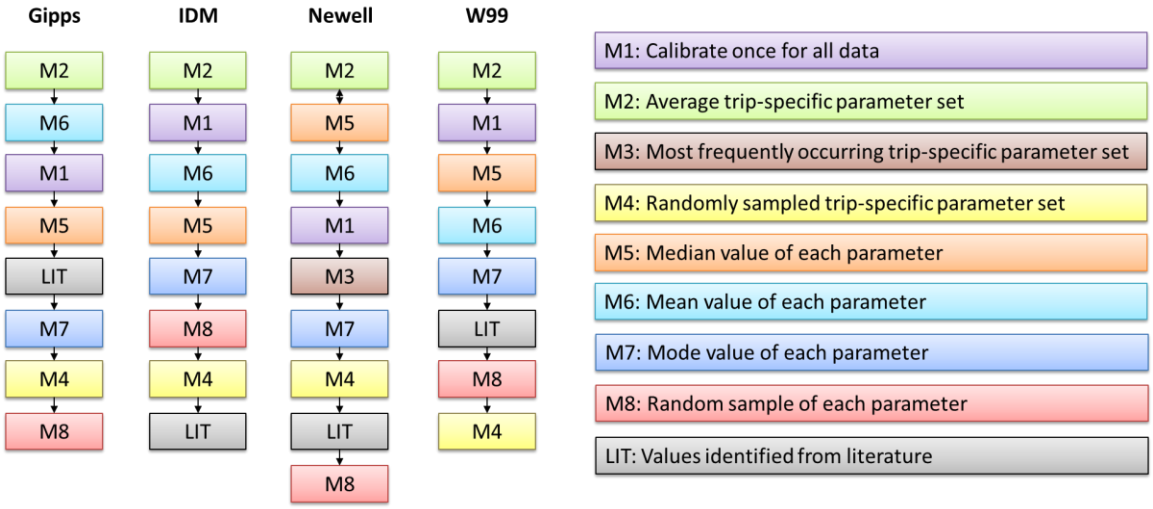


Figure 7.3 Ranking of Each Calibration Method According to the Results from the 10-Fold Cross-Validation Procedure (James, Hammit, & Boyles, 2019)

The suggested literature values for IDM performed worse than any of the eight proposed methods. For the other models, validation success of the literature values was only slightly better or slightly worse than the random sampling methods. This reinforces the importance of calibrating car-following models, rather than relying on the default values. This finding is corroborated with previous research documented in Hammit et al. (2018a).

Method 7, where the most frequently occurring parameter coefficients are selected independently, exhibited poor performance for each model. A review of the data indicates

that for at least one-third of each model's parameters, the mode was either the minimum or maximum parameter boundary. It is likely that this method performed so poorly for validation because this combination of multiple extreme parameter coefficients produced unrealistic driving behavior not replicated in naturalistic data. It should be noted that Newell was the only model that produced at least one repeated set of trip-specific calibration parameters in any of the 10 collections of calibration data (M4). For that reason, Method 3 was not applied to the other models. Moreover, the Newell representative parameter sets for Method 3 and Method 7 were identical; that is, the most frequently occurring trip-specific parameter set was equivalently the independent modes of the two parameter coefficients.

It was originally hypothesized that Method 1 would produce the lowest validation score. However, a review of calibration results showed that the calibration procedure using all following trajectories converged at a local minimum. The massive influx of car-following data (i.e., on average 3.8 hours per fold, 34 hours per calibration set) exhausted the available computational resources. With unlimited time and computational power, this method would likely produce the "best" representative parameter set; however, this method is not feasible for practical implementation.

Ultimately, there is no single best method. It is apparent from this study that methods that sample the most frequently observed behaviors (M3, M7) and methods that randomly sample parameter coefficients (M4, M8) did not perform well at producing generalizable driving behaviors. Conversely, methods that captured the average behavior (M1, M2, M5, M6) are the top performing methods for all models. The methods that captured the average driving behavior while disregarding the correlation structure between parameters (M5, M6) performed better than the methods that preserved the correlation structure between parameters but focused on capturing the most frequently observed

behavior (M3); this trend was also true for methods that randomly sampled trip-specific parameter sets (M4). Although accounting for the correlation structure is important to method performance—this is evident because M2 was the best performing method overall—these results provide evidence that capturing the average behavior is the most important factor for obtaining a representative parameter set. Because of the computational burden associated with M1 and M2, the decision regarding whether or not to preserve the underlying correlation structure (M1, M2 vs. M5, M6) depends on the available resources, project scope, data sample, and selected car-following model.

7.5. CHAPTER 7 CONCLUSIONS

Microsimulation models are an excellent resource to help agencies obtain more robust estimates of project benefits and spend their resources more effectively. However, the realism of these models is strongly dependent on the quality of the sub-models controlling individual driver behavior and the input data. The accuracy of car-following models is of particular importance to agencies interested in forecasting a facility's capacity, average travel speed, and average travel time.

Procedures for calibrating car-following models using a single driving trajectory are mature and well-documented in the literature. However, before trajectory-level data can be applied to calibrate a microsimulation model in practice, methods for identifying the most representative parameter set to describe the generalizable behavior of a group of drivers (e.g., males) or a specific driving condition (e.g., work zones) must be developed and tested. Toward this objective, this research develops eight viable methods for obtaining representative sets of calibration parameters. The methodologies are grouped into three behavioral categories: (i) average behavior, (ii) most frequently observed behavior, or (iii)

randomly sampled behavior; moreover, these methods are designed to evaluate the importance of preserving possible correlations between calibration parameters. A 100-trip sample of the SHRP2 NDS was applied to calibrate four common car-following models. A robust validation strategy from the data science literature, 10-fold cross-validation, was implemented to evaluate the performance of the eight methods.

The research findings show that the method that captured the average behavior while preserving correlations between the calibrated model parameters (M2) performed the best across all four models; this illustrates the importance of accounting for the underlying relationships between model parameters, as observed in (Kim & Mahmassani, 2011; Monteil et al., 2014). However, methods that adequately captured the average behavior while relaxing the assumption of underlying parameter correlations (M5, M6) performed better than all other methods. In other words, although the more computationally burdensome methods produce optimal results, simply taking the mean or median of the distribution of individual parameter coefficients offers a practical approach for generating a representative parameter set. For all models, these methods (M5, M6) demonstrated significantly better performance than the default parameter sets.

The availability of trajectory-level driving data is providing new opportunities to improve the accuracy of car-following models and their application in practice. The SHRP2 NDS dataset is of keen interest to modelers as it provides this high-resolution driving data for over five million trips, representing 3,400 diverse drivers in naturalistic driving conditions. Despite its potential, the sensitive nature of the collected data (i.e., personally identifiable information) and the massive quantity of trips makes it challenging to widely distribute. However, if verified calibration procedures were applied to this database, trip-specific calibration parameters for common car-following models could be disseminated without fear of privacy violations. This would offer an unprecedented amount of data to

practitioners and researchers for model development and analysis. With these data readily available, the methods presented in this chapter provide a feasible approach to improve the realism of current microsimulation practices.

Chapter 8: New Car-Following Model Calibration Framework – Using Census-Level Data for Calibration (Task 4)

8.1. MOTIVATION

In fiscal year 2016, the Federal Highway Administration (FHWA) kicked off a project funded by the Traffic Analysis and Simulation Transportation Pooled Fund Study. This project, the Transportation Systems Simulation Manual (TSSM), was intended to develop better guidance on the application of simulation to conduct transportation analyses. As part of this project, FHWA hosted a series of stakeholder meetings and virtual roundtables to better assess the current climate of applying simulation to evaluate transportation alternatives and understand needs and gaps within the community. From this assessment, it was revealed that calibration of complex traffic analysis models is the most commonly cited challenge facing State Departments of Transportation (DOTs). These conversations also found that a lack of robust data and requests for increasingly complex traffic analyses, without additional funding to cover incurred costs, have pushed practitioners to cut costs by cutting corners, such as omitting the calibration of the driver behavior component of the simulation model and relying on default model parameters instead.

This dissertation explores a new framework for driver behavior model calibration that aims to reduce the burden of calibrating driver behavior in microsimulation models. The assumptions in the framework are as follows. First, this framework assumes that high-resolution data describing driver behavior, such as the second Strategic Highway Research Program (SHRP2) Naturalistic Study Data (NDS), are available. This framework also assumes that these data have been processed (see Section 3.1) and used to obtain trip-specific calibration coefficients (see Section 3.2) for each available trajectory. This is a reasonable assumption because eventually trip-specific calibration coefficients for the

SHRP2 NDS trips could be posted freely online for practitioner use. Some trip statistics for the SHRP2 NDS trips are currently available freely and openly on the SHRP2 Insight website; although calibrated car-following model parameter values are not currently posted, it is not unreasonable to assume that this information could one day be made available freely online given that it is not protected personally identifiable information (PII).

This new framework hypothesizes that inter-driver heterogeneity is explainable using census-level details about the driver (i.e., demographics data). This framework takes advantage of existing large-scale trajectory-level data collection efforts, such as the SHRP2 NDS, to pre-process and calibrate car-following models and shift the calibration burden toward the back-end of the effort, significantly increasing the practicality of accounting for inter-driver differences in microsimulation models. The proposed framework for using census-level data in the calibration process is shown in Figure 8.1.

The remainder of this chapter is organized as follows. Section 8.2 briefly overviews the methodology applied in this chapter. Section 8.3 covers the results of the chapter. Finally, Section 8.4 discusses the conclusions, limitations, and planned future research inspired by this effort.

It should be noted that this chapter is not able to determine the accuracy of using this framework for car-following model calibration, compared against traditional methods of car-following model calibration, as the simulation network is hypothetical and ground-truth verification data is not available. However, this chapter does show that different proportions of driver groups do produce considerably different key performance metrics of a microsimulation model (e.g., capacity), which could impact the results of an alternatives assessment and ultimately influence which projects are selected for funding. At a minimum, this chapter successfully provides evidence suggesting that the influence of

inter-driver differences on key transportation performance metrics are worthy of additional analysis.

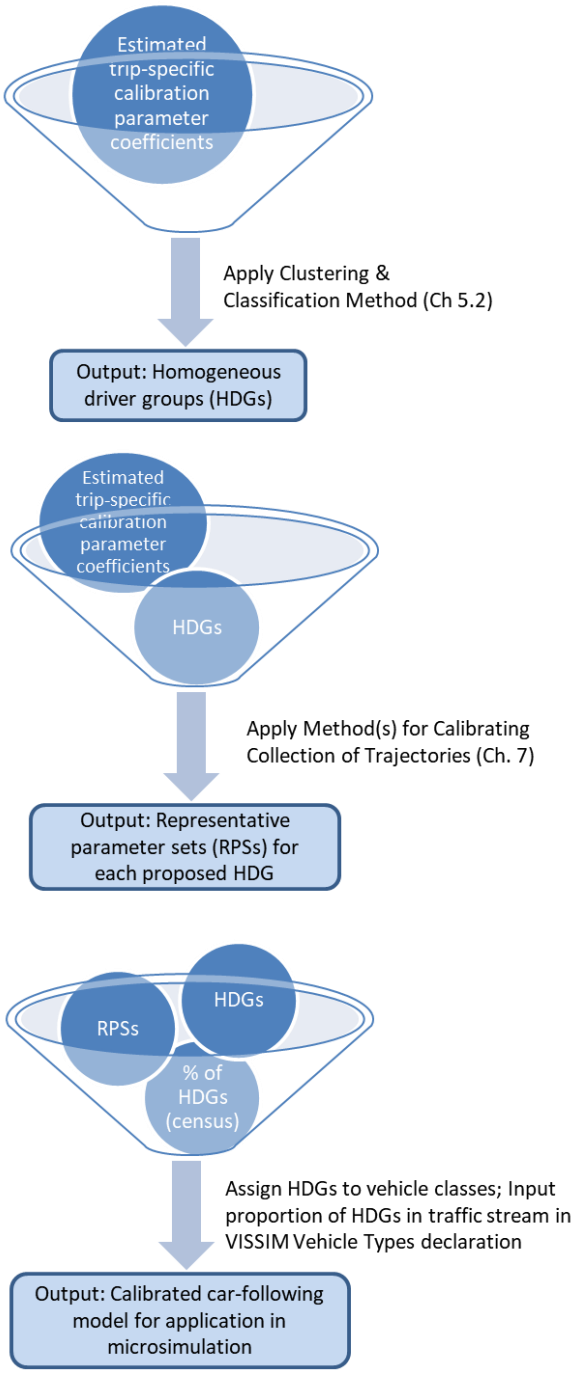


Figure 8.1 Proposed Car-Following Model Calibration Framework

8.2. METHODOLOGY

Estimated car-following model calibration coefficients are required for each trajectory as input to the framework. The methodologies adopted by this dissertation for acquiring trip-specific calibration coefficients are all summarized in Chapter 3. For the sake of brevity, the reader is referred to Section 3.1.3 for an overview of the Wyoming DOT sample of the SHRP2 NDS dataset used for this analysis, Section 3.2.3 for an overview of three car-following models applied in this chapter (i.e., Gipps, Intelligent Driver Model (IDM), and Wiedemann 99 (W99)), and Section 3.2.4 for an overview of the nonlinear optimization problem and genetic algorithm used to identify the optimal car-following model calibration parameters for each trip in the 665-trip sample available for analysis.

Identifying Homogeneous Driver Groups

In Section 5.3 of this dissertation, a data-driven methodology for identifying clusters of similar driving behavior, or homogeneous driver groups, was discussed. To illustrate the validity of this method, Section 5.3 initially applies the clustering and classification methodology to each calibration parameter, individually. This allowed for visual inspection of the developed clusters. In this chapter, homogeneous driver groups were developed considering all calibration parameters belonging to a car-following model, holistically. The *Expectation Maximization* clustering algorithm was applied (see Section 5.3.1 for algorithm specifics). The *Expectation Maximization* algorithm identified that four clusters was the optimal number of clusters of driving behavior for the Gipps, IDM, and W99 car-following models, separately; this algorithm also assigned each trip to their proper cluster.

Next, classification algorithms were used to identify which driver attributes can be used to accurately assign a driver to their cluster ID. The *ZeroR*, *OneR*, *PART Decision Rules*, and *J48 Decision Tree* classification algorithms were considered; for additional details on the four classification algorithms, see Section 5.3.1. This procedure found that the *PART Decision Rules* algorithm best classified the Gipps homogeneous driver group clusters, the *OneR* algorithm best classified the IDM homogeneous driver group clusters, and the *J48 Decision Tree* algorithm best classified the W99 homogeneous driver group clusters. Marital status and age group were the most predictive attributes of homogeneous driver groups. The driver demographics used to divide the 665 trips into their homogeneous driver groups for simulation by the Gipps, IDM, and W99 car-following models are shown in Table 8.1, Table 8.2, and Table 8.3, respectively.

Table 8.1 Gipps Homogeneous Driver Groups Developed using *PART Decision Rules* Algorithm*

Cluster 1	17. Income = 110–149k & Age Group = 35–39 21. Age Group = 35–39 23. Any remaining drivers
Cluster 2	3. Age Group = 25–29 & Gender = F & Income = 50–69k 6. Marital Status = divorced & Age Group = 35–39 13. Income = 110–149k & Age Group = 40–44 15. Income = 50–69k 20. Education = College degree 22. Gender = F
Cluster 3	1. Marital Status = unmarried partners 5. Age Group = 25–29 9. Education = No college degree & Age Group = 70+* & Income = 40–49k 10. Education = No college degree
Cluster 4	2. Age Group = 25–29 & Education = College degree & Marital Status = single 4. Age Group = 60–69* 7. Income = 50–69k & Education = College degree 8. Education = No college degree & Age Group = 70+* & Income = Under 39k* & HH Size* = 1 11. Income = 70–99k 12. Marital Status = single 14. Income = 110–149k & Marital Status = married & Education = Graduate Degree 16. Marital Status = married & Income = 110–149k & Age Group = 45–59* 18. Income = 150k+ & Age Group = 70+* 19. Income = 150k+ & Gender = M & Age Group = 30–34 & HH Size* = c: 3

* Given the nature of decision rules, to correctly classify the drivers, the data that is covered by each rule must be removed from the master dataset in the order that the rules are numbered in the table.

Table 8.2 IDM Homogeneous Driver Groups Developed Using *OneR* Algorithm

Cluster 1	divorced single
Cluster 2	married
Cluster 3	widow(er)
Cluster 4	unmarried partners

Table 8.3 Wiedemann 99 Homogeneous Driver Groups Developed Using *J48 Decision Tree* Algorithm

Cluster 1	Marital Status = single & Age Group = 25–29 & Income = 40–49k Marital Status = single & Age Group = 25–29 & Income = Under 39k Marital Status = single & Age Group = 20–24 Marital Status = unmarried partners
Cluster 2	Marital Status = divorced Marital Status = single & Age Group = 45–59 Marital Status = single & Age Group = 40–44 Marital Status = married
Cluster 3	Marital Status = widow(er)
Cluster 4	Marital Status = single & Age Group = 35–39 Marital Status = single & Age Group = 30–34

Obtaining Calibration Parameters to Describe Collections of Trajectories

In Chapter 7 of this dissertation, methods for identifying representative parameter sets to describe a collection of trajectories were evaluated and validated. From this analysis, it was determined that the best method varied according to the car-following model applied. For the Gipps and IDM car-following models, simply taking the average of each calibration parameter for the collection of trajectories, independently, proved practical and sufficiently accurate; for the W99 car-following model, taking the median of each calibration parameter for the collection of trajectories, independently, was found to work best from a practical perspective. Thus, once all of the trips are segmented into their homogeneous driver groups according to the definitions in Table 8.1, Table 8.2, and Table 8.3, the representative parameter sets for the Gipps, IDM, and W99 models are identified using the most well-

performing, yet practical, methods, as identified in Chapter 7. This produced estimated sets of calibration parameters for each model, which are documented in Table 8.4, Table 8.5, and Table 8.6.

Table 8.4 Representative Calibration Parameter Sets Variation Across Gipps Homogeneous Driver Groups

Model Parameters		**Default**	**C1**	**C2**	**C3**	**C4**
Gipps	Desired Velocity [m/s]	35.0	33.0	33.4	33.4	31.7
	Desired Acceleration [m/s^2]	2.0	1.4	1.1	1.2	1.5
	Reaction Time [s]	0.7	1.3	0.4	0.7	0.8
	Desired Deceleration [m/s^2]	-3.0	-3.0	-3.0	-3.1	-2.4
	Predicted Maximum Deceleration of Lead Vehicle [m/s^2]	-3.5	-2.5	-2.9	-2.8	-2.2
	Minimum Standstill Distance [m]	1.0	4.9	3.0	5.0	4.7

Some initial reactions to the magnitude of the representative parameter sets for the Gipps homogeneous driver groups are discussed next. Table 8.4 shows that the default Gipps desired velocity parameter is significantly higher than the calibrated desired velocity parameter across all four homogeneous driver groups. However, the default desired acceleration parameter is only marginally larger than the calibrated desired acceleration parameter for the four clusters of similar driving behavior. The default reaction time

parameter is contained within the range of calibrated reaction time parameters. The default and calibrated desired deceleration parameters are very close in magnitude, but the default maximum perceived deceleration of the leading vehicle is considerably higher than the calibrated values for the four homogeneous driver groups; this is an interesting observation because correlation analyses have shown that these two parameters are highly correlated (Hammit et al., 2018). Moreover, with the default parameters, the desired deceleration estimate is smaller than the estimated perceived deceleration of the leading vehicle. However, for all four clusters of calibrated driving behavior, the desired acceleration estimate is larger than the perceived deceleration of the leading vehicle; this can lead to instability of the car-following behavior and is somewhat of a cause for concern. Finally, the default estimate for the minimum standstill distance is significantly lower than the calibrated values for the minimum standstill distance parameter, with perhaps the exception of Cluster 2. As observed in previous chapters, this is likely an artifact of the data (i.e., collected on freeways where not all trips experienced congested conditions).

Table 8.5 Representative Calibration Parameter Sets Variation Across IDM Homogeneous Driver Groups

Model Parameters		Default	C1	C2	C3	C4
Intelligent Driver Model	Desired Velocity [m/s]	35.0	33.1	33.4	29.4	35.9
	Free Acceleration Component	4.0	54.3	36.7	24.6	42.6
	Desired Time Gap [s]	1.5	0.7	0.7	1.7	0.5
	Jam Distance [m]	2.0	4.0	3.4	5.3	4.0
	Desired Acceleration [m/s^2]	1.4	1.1	0.8	0.3	0.9
	Desired Deceleration [m/s^2]	2.0	2.0	2.7	2.4	3.3

The IDM representative parameter sets for the four homogeneous driver groups are shown in Table 8.5. Cluster 4, which represents unmarried partners, was estimated to have the highest desired velocity parameter value; the smallest desired velocity parameter was observed for Cluster 3, or the widow(er) subcategory. The calibrated free acceleration components are all significantly larger than the default value. The default desired time gap parameter is within the range of the calibrated desired time gap parameters, although it is on the higher end of the spectrum. The smallest desired time gap parameter was estimated for Cluster 4, or the unmarried partners drivers, while the largest was estimated for Cluster 3, the widow(er) subcategory. The smallest jam distance belongs to the default parameter set. The largest jam distance was estimated for Cluster 3, or the widow(er) drivers. The largest desired acceleration parameter value occurred in the default parameter set. The

smallest desired acceleration parameter value was estimated for Cluster 3, the widow(er) subgroup. The smallest desired deceleration parameter was estimated for Cluster 1, which represents both single and divorced drivers; this is equivalent to the default desired deceleration parameter value. The largest desired deceleration parameter was estimated for Cluster 4, or the unmarried partners subgroup of drivers.

Table 8.6 Representative Calibration Parameter Sets Variation Across W99 Homogeneous Driver Groups

Model Parameters		Default	C1	C2	C3	C4
W99	CC0: Standstill Distance [m]	1.5	5.1	3.6	6.0	4.5
	CC1: Spacing Time [s]	1.3	0.6	0.6	1.0	0.5
	CC2: Following Variation, Max Drift [m]	4.0	13.8	11.9	12.4	13.2
	CC3: Threshold for Entering Following [s]	-12	-25.6	-23.8	-23.8	-24.1
	CC4: Negative Following Threshold [m/s]	-0.25	0.0	-0.1	0.0	-0.1
	CC5: Positive Following Threshold [m/s]	0.35	0.7	1.0	1.0	1.2
	CC6: Speed Dependency of Oscillation [10^{-4} rad/s]	0.0006	1.3	2.1	1.4	1.8
	CC7: Oscillation Acceleration [m/s^2]	0.25	1.3	0.6	1.1	1.0
	CC8: Standstill Acceleration [m/s^2]	2.0	1.3	1.6	1.5	1.7
	CC9: Acceleration at 80kph [m/s^2]	1.5	0.1	0.1	0.1	0.2
	Desired Velocity [m/s]	35.0	34.2	33.9	31.6	31.9

The W99 representative parameter sets for the four homogeneous driver groups are shown in Table 8.6. The smallest standstill distance parameter value was associated with the default model parameters. The smallest calibrated standstill distance (CC0) was estimated for Cluster 2, which is comprised of divorced, married, and single drivers in their 40s and 50s. The largest standstill distance was estimated for Cluster 3, which was comprised of widow(er) drivers. The largest spacing time (CC1) coefficient was associated with the default model parameters. The smallest spacing time (CC1) parameter was estimated for Cluster 4, which was comprised of single drivers in their 30s. The largest calibrated spacing time (CC1) coefficient was estimated for Cluster 3, which was comprised of widow(er) drivers. The default oscillation acceleration (CC7) parameter value was significantly smaller than the calibrated parameter values. The smallest oscillation acceleration (CC7) parameter value was estimated for Cluster 2, which is comprised of divorced, married, and single drivers in their 40s and 50s. The largest oscillation acceleration (CC7) coefficient was estimated for Cluster 1, which was comprised of single drivers under the age of 30 and unmarried partners. The default standstill acceleration (CC8) parameter value was larger than the calibrated values. The smallest standstill acceleration (CC8) parameter value was estimated for Cluster 1, while the largest value was estimated for Cluster 4. Finally, the default desired velocity parameter value is significantly higher than the calibrated values. The smallest desired velocity parameter value was calibrated for Cluster 3, while the largest value was estimated for Cluster 1.

Microsimulation Model Development

This dissertation uses a recommended simple weaving segment from the Sixth Edition of the Highway Capacity Manual (Transportation Research Board, 2010), which

includes four-lanes in one-direction with three on- and off-ramps, as shown in Figure 8.2. This network was built in PTV VISSIM 9.

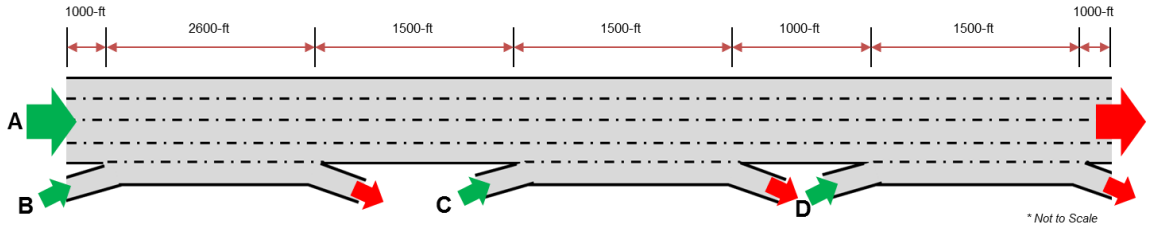


Figure 8.2 Simple Freeway Weaving Section Modeled in VISSIM (Hammit et al., 2019)

The Gipps and IDM driver behavior logic are not included in the PTV VISSIM software. Thus, to simulate vehicles following this logic, the External Driver Model Dynamic Linking Library (DLL) Interface of VISSIM was employed. The driving behavior logic of VISSIM can be overwritten using a DLL, written in C/C++, that is assigned to a vehicle type in the VISSIM software. For each timestep, VISSIM calls the DLL code for each vehicle whose type is assigned to the DLL. VISSIM passes the current state of the vehicle and its surrounding vehicles to the DLL; the available vehicle attributes that can be passed are noted in the C++ header file. The DLL computes the acceleration/deceleration and passes the updated state of the vehicle back to VISSIM for visualization.

Three source codes are provided in the VISSIM installation package for DLL creation:

- DriverModel.h, which is used to specify which attributes can be passed between the DLL and VISSIM.
- DriverModel,cpp, which is the main source file of the DLL. This is where the external logic is coded. Each DriverModel.cpp code must contain and export three functions, which are called from VISSIM:

DriverModelSetValue, DriverModelGetValue, and DriverModelExecuteCommand.

- DriverModel.vcxproj, which is the Visual C++ project file for the driver model DLL.

The DriverModel.cpp file is what is edited to contain the external vehicle behavior logic. In the DriverModelSetValue function (see Figure 8.3), the modeler can determine which attributes they want to retrieve from VISSIM about the subject vehicle and its surrounding vehicles. The DriverModel.h file is helpful for determining which attributes (e.g., DRIVER_DATA_VEH_ID, DRIVER_DATA_VEH_VELOCITY, AND DRIVER_DATA_NVEH_REL_VELOCITY) are of interest to the modeler. The DriverModelSetValue function can retrieve up to a five-by-five matrix of information about surrounding vehicles. In this matrix, the current vehicle is vehicle [2][2]; thus, to obtain information about the leading vehicle, the modeler is interested in retrieving information about vehicle [2][3].

The DriverModelGetValue function (see Figure 8.4) allows the modeler to pass manipulated information about the vehicle (i.e., computed desired acceleration using external logic) back to VISSIM. The attributes that can be passed back to VISSIM are defined in the DriverModel.h file.

Finally, in the DriverModelExecuteCommand function (see Figure 8.5), the external driver behavior logic is programmed with the appropriate syntax. The ControlVehicle function (see Figure 8.6) within the DriverModelExecuteCommand function is where the vehicle control logic is housed. The logic featured in Figure 8.6 is the IDM car-following model. This logic was replaced with the Gipps and W99 car-following models according to the algorithms in Section 3.2.3 when developing the Gipps and W99

Driver Model DLLs. A separate DriverModel.cpp file was developed for every homogeneous driver group.

```
DRIVERMODEL_API  int  DriverModelSetValue(long   type,
    long   index1,
    long   index2,
    long   long_value,
    double double_value,
    char   *string_value)
{
    /* Sets the value of a data object of type <type>, selected by <index1> */
    /* and possibly <index2>, to <long_value>, <double_value> or            */
    /* <*string_value> (object and value selection depending on <type>).    */
    /* Return value is 1 on success, otherwise 0.                           */

    switch (type) {
    case DRIVER_DATA_VEH_ID:
        current_veh = long_value;
        return 1;
    case DRIVER_DATA_VEH_VELOCITY:
        from_VISSIM_spd_vehicle = double_value;
        return 1;
    case DRIVER_DATA_VEH_ACCELERATION:
        from_VISSIM_acc_vehicle = double_value;
        return 1;
    case DRIVER_DATA_VEH_LENGTH:
        Leng_vehicle = double_value;
        return 1;
    case DRIVER_DATA_VEH_DESIRED_VELOCITY:
        from_VISSIM_desired_velocity = double_value;
        return 1;
    case DRIVER_DATA_NVEH_ID:
        AdjVehicles[index1 + 2][index2 + 2] = long_value;
        return 1;
    case DRIVER_DATA_NVEH_DISTANCE:
        AdjVehicleSpaceHeadway[index1 + 2][index2 + 2] = double_value;
        return 1;
    case DRIVER_DATA_NVEH_REL_VELOCITY:
        AdjVehiclesSpeed[index1 + 2][index2 + 2] = from_VISSIM_spd_vehicle - double_value;
        return 1;
    case DRIVER_DATA_NVEH_ACCELERATION:
        AdjVehiclesAcc[index1 + 2][index2 + 2] = double_value;
        return 1;
    case DRIVER_DATA_NVEH_LENGTH:
        AdjVehiclesLength[index1 + 2][index2 + 2] = double_value; // revised for VISSIM 7.
        return 1;
    default:
        return 0;
    }
```

Figure 8.3 Screen Capture of DriverModelSetValue Function

```
DRIVERMODEL_API  int  DriverModelGetValue(long   type,
    long   index1,
    long   index2,
    long   *long_value,
    double *double_value,
    char   **string_value)
{
    /* Gets the value of a data object of type <type>, selected by <index1> */
    /* and possibly <index2>, and writes that value to <*double_value>,     */
    /* <*float_value> or <**string_value> (object and value selection       */
    /* depending on <type>).                                                 */
    /* Return value is 1 on success, otherwise 0.                            */

    switch (type) {

    case DRIVER_DATA_DESIRED_ACCELERATION:
        *double_value = desired_acceleration;
        return 1;
    default:
        return 0;
    }
}
```

Figure 8.4 Screen Capture of DriverModelGetValue Function

```
DRIVERMODEL_API  int  DriverModelExecuteCommand(long number)
{
    /* Executes the command <number> if that is available in the driver */
    /* module. Return value is 1 on success, otherwise 0.               */

    switch (number) {
    case DRIVER_COMMAND_INIT:
        now = time(0);
        dt = ctime(&now);
        return 1;
    case DRIVER_COMMAND_CREATE_DRIVER:
        VehTargetLane[current_veh] = 0;
        return 1;
    case DRIVER_COMMAND_KILL_DRIVER:
        VehTargetLane.erase(current_veh);
        return 1;
    case DRIVER_COMMAND_MOVE_DRIVER:
        leaderID = AdjVehicles[2][3];
        followerID = AdjVehicles[2][1];
        front_dist = AdjVehicleSpaceHeadway[2][3];
        rear_dist = -AdjVehicleSpaceHeadway[2][1];

        ControlVehicle();
        return 1;
    default:
        return 0;

            }

}
```

Figure 8.5 Screen Capture of DriverModelExecuteCommand Function

```
int ControlVehicle()
{
    // Values from Simulation
    double relative_spaceheadway = AdjVehicleSpaceHeadway[2][3]; // [m]
    double LeadVeh_Length = AdjVehiclesLength[2][3]; // [m]
    double relative_spacegap = relative_spaceheadway - LeadVeh_Length; // [m]
    double LeadVeh_Spd = AdjVehiclesSpeed[2][3];// [m/s]
    double LeadVeh_ID = AdjVehicles[2][3]; // [unitless]
    double relative_speed = from_VISSIM_spd_vehicle - LeadVeh_Spd; // follower - leader [m/s]
    double FollowVeh_Speed = from_VISSIM_spd_vehicle; // [m/s]
    double FollowVeh_Acc = from_VISSIM_acc_vehicle; //[m/s^2]
    double FollowVeh_ID = current_veh;
    double LeadVeh_LaneID = AdjVehiclesCurrentLane[2][3];
    double FollowVeh_LaneID = current_lane;
    // Calculated Values
    double a_IDM = 0.0;
    double a_free = 0.0;
    double a_brake = 0.0;
    double s_star = 0.0; // desired safe gap
    double velocityratio = 0.0;
    double spacingratio = 0.0;
    // User Defined Parameters
    double v_0 = 33.1; // desired speed in m/s
    double delta = 54.3; // free acceleration compontent [unitless]
    double T = 0.7; // desired time gap [s]
    double a = 1.1; // maximum desired acceleration [m/s2]
    double b = 2.2; // maximum desired deceleration [m/s2]
    double s_0 = 4; // jam distance [m]
    desired_velocity = v_0;

    if (LeadVeh_ID != -1)
    {

        s_star = s_0 + max(0.0, (FollowVeh_Speed * T) + ((FollowVeh_Speed * relative_speed) / (2 * sqrt(a * b))));
        spacingratio = s_star / relative_spacegap;
        velocityratio = FollowVeh_Speed / v_0;
        a_free = a * (1 - pow(velocityratio, delta));
        a_brake = (-a) * spacingratio * spacingratio;
        a_IDM = a_free + a_brake;
        desired_acceleration = a_IDM;

    }

    return 1;

}
```

Figure 8.6 Screen Capture of ControlVehicle Function featuring IDM Car-Following Model

Data collection points were designated at five points throughout the network featured in Figure 8.2. The data is aggregated every 20s and converted to hourly flow rates. Using the fundamental relationship between space-mean speed, volume, and density, density is computed.

The simulated traffic flow data for varying proportions of driver groups is shown on flow-density, speed-flow, and speed-density fundamental diagrams and discussed in the

next section. For the purpose of this case study, five sets of simulation runs were completed for each car-following model:

- 100% of vehicle types assigned to Cluster 1;
- 100% of vehicle types assigned to Cluster 2;
- 100% of vehicle types assigned to Cluster 3;
- 100% of vehicle types assigned to Cluster 4; and
- Equal proportions (i.e., 25%) of all homogeneous driver groups.

8.3. CASE STUDY RESULTS

This section is separated by car-following model logic. The simulated Gipps results are shared in Figure 8.7, Figure 8.8, Figure 8.9, and Table 8.7. The simulated IDM results are shared in Figure 8.10, Figure 8.11, Figure 8.12, and Table 8.8. Finally, the simulated W99 results are documented in Figure 8.13, Figure 8.14, Figure 8.15, and Table 8.9.

Gipps Car-Following Model Results

The flow-density fundamental diagrams produced using the Gipps car-following model are shown in Figure 8.7. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. All fundamental diagrams produced with calibrated model parameters (i.e., Figure 8.7b–Figure 8.7f) produce evidence of the capacity drop phenomenon; this is not as strongly suggested by the fundamental diagram produced using the default Gipps parameters (i.e., Figure 8.7a). Moreover, Figure 8.7b, which was simulated with equal proportions of the four homogeneous driver groups, has features from the four fundamental diagrams simulated with 100% of each of the

homogeneous groups, independently (i.e., Figure 8.7c through Figure 8.7f); this suggests that by simulating proportions of driver groups, a modeler is able to better capture diverse driving behaviors. For all fundamental diagrams in Figure 8.7, there is a strong linear trend between volume and density in uncongested conditions, as evident by the colored squares on the left side of the graphs; however, in congested conditions, there is much more variation in the data with few obvious trends. Interestingly, Cluster 2's fundamental diagram (Figure 8.7d) most closely resembles the trends observed in the fundamental diagram simulated with default parameters (Figure 8.7a); it is not immediately obvious what causes this similarity. Finally, the highest flow rate was observed at a relatively lower density with Cluster 1's optimal parameters; this suggests a fast transition from uncongested to congested conditions occurred with the parameter set.

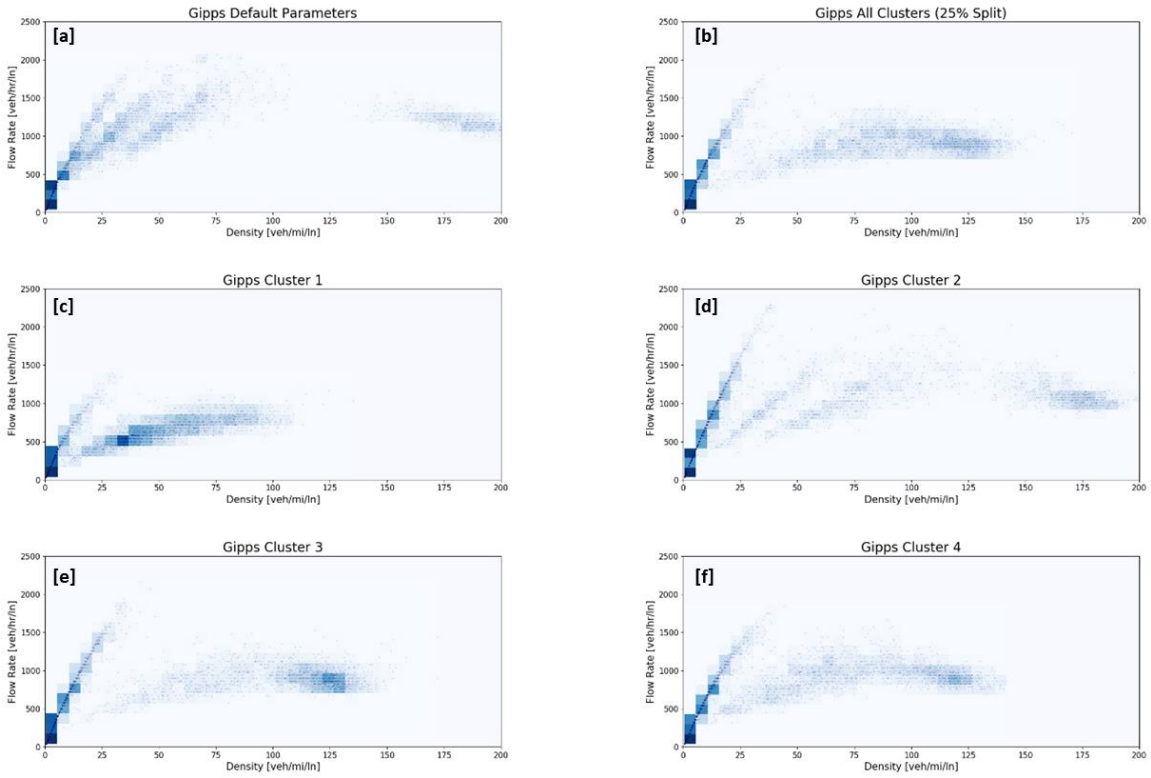


Figure 8.7 Flow-Density Fundamental Diagrams for Different Proportions of the Gipps Homogeneous Driver Groups

The speed-flow fundamental diagrams produced using the Gipps car-following model are shown in Figure 8.8. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. All fundamental diagrams produced using the Gipps car-following model, shown in Figure 8.8b through Figure 8.8f, have a large number of observations of high-speed/low-volume and low-speed/moderate-volume conditions. Visual observations of traffic flow revealed that this is because traffic flow deteriorates very quickly from uncongested to congested conditions, without much of a transitional period. The fundamental diagrams produced using the default parameters, shown in Figure 8.8a, has better representation of the spectrum of speeds (i.e., from zero

to free-flow speed) compared to traffic simulated with the other calibrated parameter sets. This suggests that there is less instability in the flow of traffic using these model parameters, which is consistent with observations in Section 8.2; additionally, the maximum flow rate for this fundamental diagram was observed at the lowest speed, further supporting this hypothesis. The traffic flow produced using Cluster 2's optimal parameter set, shown in Figure 8.8b, observed its maximum flow rate at a much higher speed compared to the other clusters; this is an indicator of more instability in queue discharge rates.

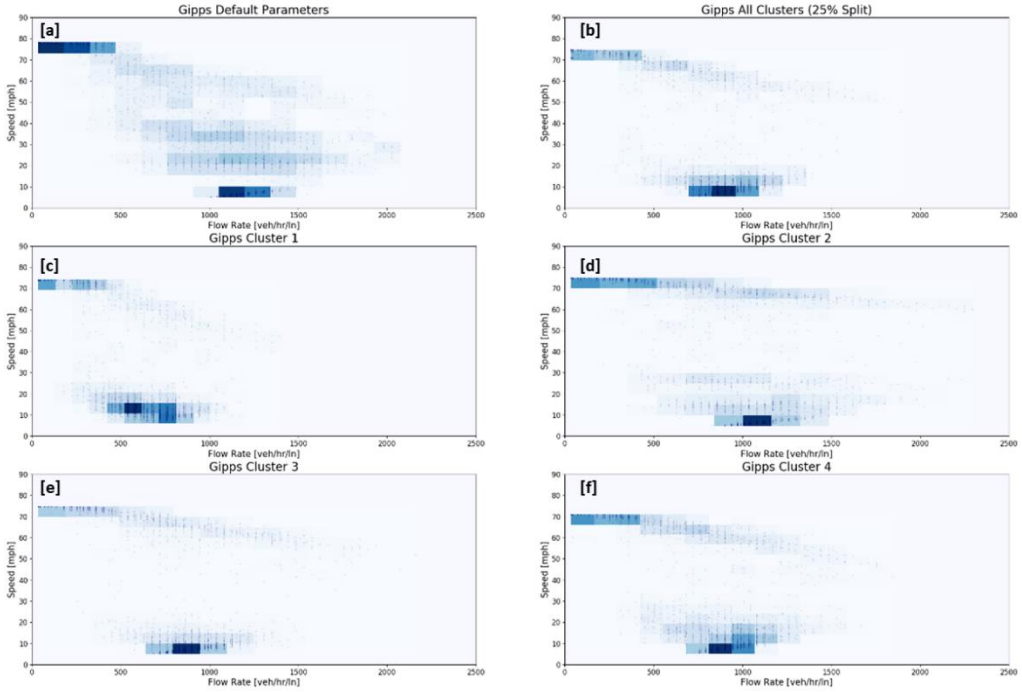


Figure 8.8 Speed-Flow Fundamental Diagrams for Different Proportions of the Gipps Homogeneous Driver Groups

The speed-density fundamental diagrams produced using the Gipps car-following model are shown in Figure 8.9. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed

data points in that region of the fundamental diagram. Visual observation of Figure 8.9 suggests that the Gipps model does not produce a linear relationship between speed and density. Instead, traffic flow deterioration appears to follow a negative exponential trend; this is consistent with Underwood's hypothesis of the relationship between speed and density.

The Gipps model tends to produce higher jam densities than the W99 and IDM models, for both the default and calibrated parameters (see Figure 8.9a through Figure 8.9f). This is likely because the Gipps model, programmed using the External Driver Model Interface, was not fully collision free on a multilane facility. Gipps was the only model with a physical reaction time (i.e., the model stores a desired velocity until the reaction time passes and the driver can react to, now outdated, information). In cases of lane changes, it takes the driver the reaction time to realize that it is now following a different driver with a different velocity (i.e., in the case of Cluster 1, this took 13-time steps). This infrequently produced instances where the driver passed through the leading vehicle and likely resulted in overestimated jam density values. The default parameters produced the highest jam density (see Figure 8.9a). The second highest jam density was observed for Cluster 2 (see Figure 8.9d). It should be noted that these two parameter sets were observed to have the smallest minimum standstill distance.

In accordance with Greenshield's theory, linear regression models between speed and density were calculated for each of the fundamental diagrams. Cluster 1 was found to have the steepest slope (see Figure 8.9c); this suggests traffic deteriorates more quickly from free-flow speed to jam density. The default parameter set produced the most gradual slope. These observations are consistent with other default parameter fundamental diagram conclusions (see Figure 8.8).

The default parameters produced the smallest goodness-of-fit metric for the linear regression equation; this suggests the least agreement with Greenshield's hypothesis of a linear relationship between speed and density. Cluster 3 produced the largest goodness-of-fit metric, which assumes there exists a linear relationship between speed and density (see Figure 8.9e).

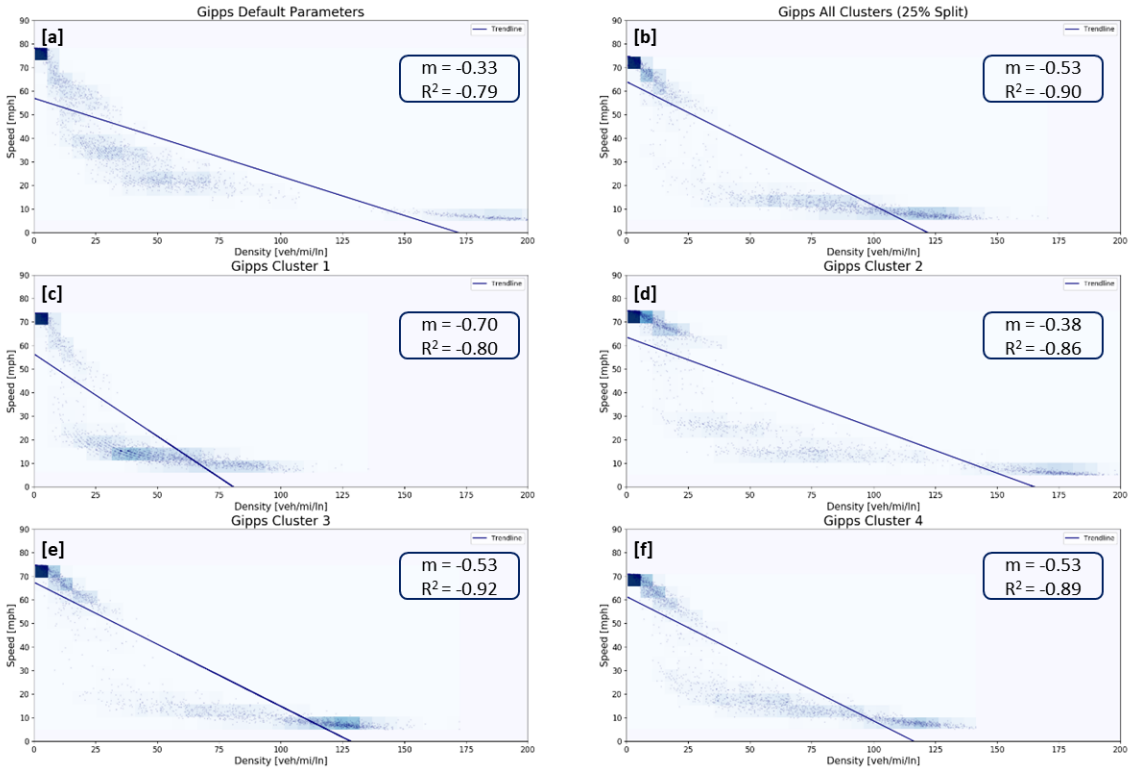


Figure 8.9 Speed-Density Fundamental Diagrams for Different Proportions of the Gipps Homogeneous Driver Groups

Finally, Table 8.7 shows the jam density and capacities of the network (see Figure 8.2) simulated using the Gipps car-following behavior according to the different representative parameter sets for the four homogeneous driver groups. Traffic flow simulated with default parameter values achieve a significantly higher jam density. The

only parameter set that produced a comparable jam density is Cluster 2; as previously documented, these two parameter sets also share the smallest minimum standstill distance parameter. The smallest jam density was observed for Cluster 1.

The capacity results are a bit mixed. All models produced reasonable estimates for capacity. The smallest capacity was produced by Cluster 1's parameter set, while the largest capacity was produced by Cluster 2's parameter set. There is significant variation in network capacity as a function of calibration parameter sets used across the five simulations conducted (i.e., a 900 veh/hr/ln difference), underscoring the importance of proper calibration (i.e., these parameters significantly impact the outputs of these models).

Table 8.7 Variation in Jam Density and Capacity Across Model Parameters (Gipps)

	Jam Density (veh/h)	**Change from Default (%)**	**Capacity (veh/mi)**	**Change from Default (%)**
Gipps Default Parameters	171	--	2070	--
Gipps All Clusters (25% Split)	121	-29%	1890	-9%
Gipps Cluster 1	80	-53%	1395	-33%
Gipps Cluster 2	165	-4%	2295	11%
Gipps Cluster 3	128	-25%	2160	4%
Gipps Cluster 4	116	-32%	1845	-11%

Intelligent Driver Model Car-Following Model Results

The flow-density fundamental diagrams produced using the IDM car-following model are shown in Figure 8.10. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. For all fundamental diagrams (Figure

8.10a through Figure 8.10f), there is a strong linear trend between volume and density in uncongested conditions, as evident by the colored squares on the left side of the graphs; however, in congested conditions, there is much more variation in the data with few obvious trends. The exception is for widow(er) drivers, which were represented by Cluster 3's fundamental diagram (see Figure 8.10c). Cluster 3's parameters produced the most well-defined traffic flow patterns in congested conditions (i.e., the most colored squares, which indicates a higher frequency of points). Overall, the traffic flow deterioration trend produced by the IDM car-following model seems a bit less sporadic than what was produced by the Gipps car-following model. However, as evident in Figure 8.10a, the default parameter produced the most varied observations (i.e., the fewest colored squares) in congested conditions.

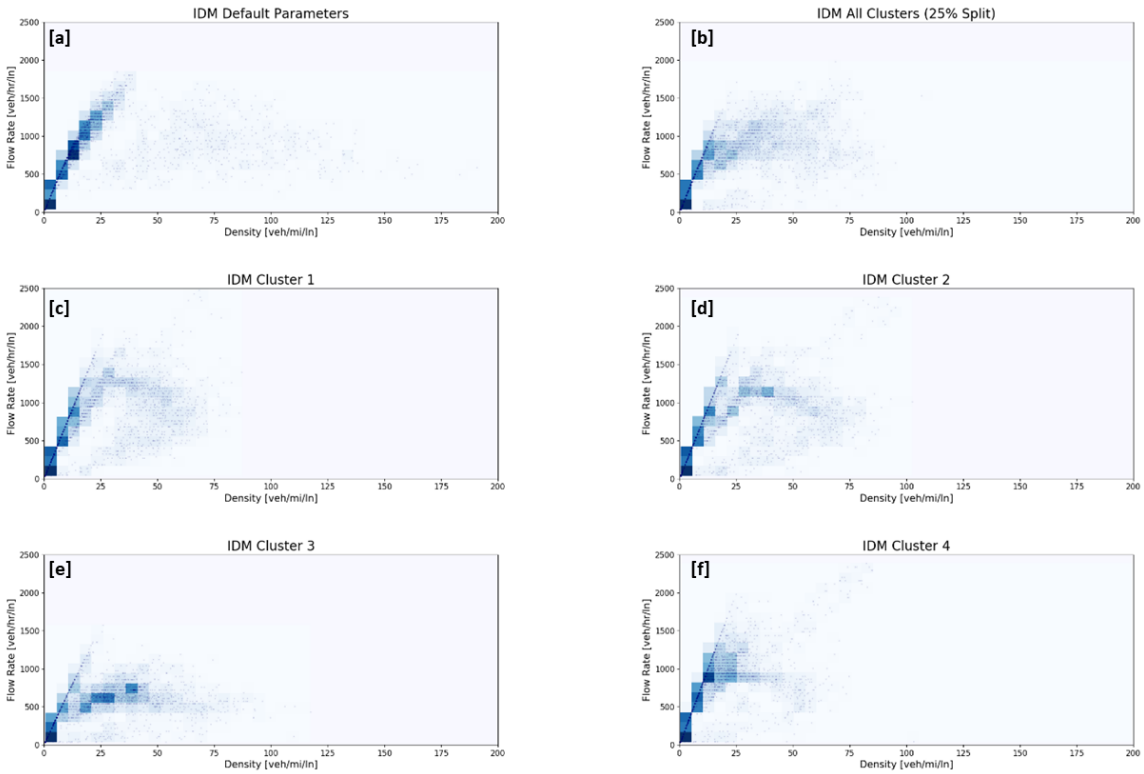


Figure 8.10 Flow-Density Fundamental Diagrams for Different Proportions of the IDM Homogeneous Driver Groups

The speed-flow fundamental diagrams produced using the IDM car-following model are shown in Figure 8.11. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. As suggested in Figure 8.11, drivers following according to Cluster 2's representative parameter coefficients were able to maintain their desired speed for longer; this is evident by the darker colored squares associated with higher speeds and higher volumes. Behaviorally speaking, this suggests drivers are less impacted by surrounding drivers because of these car-following model calibration coefficients.

The fundamental diagrams produced by the default parameter coefficients (see Figure 8.11a) and Cluster 3's representative parameter coefficients (see Figure 8.11e) observe their maximum flow rates at relatively higher speeds; this indicates there is greater instability in the queue discharge (i.e., flow between under saturated and oversaturated conditions). This is visually evident in Figure 8.11 because of the gap in the data at moderate speeds observed for these two parameter sets, relative to other specifications.

Patterns from the Cluster 1, Cluster 2, Cluster 3, and Cluster 4 fundamental diagrams (Figure 8.11c through Figure 8.11f) are all represented in the fundamental diagram constructed using equal proportions of clusters (Figure 8.11b); this suggests that by simulating proportions of driver groups, a modeler is able to better capture diverse driving behaviors. The flow breakdown from uncongested to congested flow is also more easily observable than when default parameters are used (i.e., no significant gaps in the data at moderate speeds). This underscores the value of using multiple parameter sets to simulate diverse driving behavior in naturalistic data.

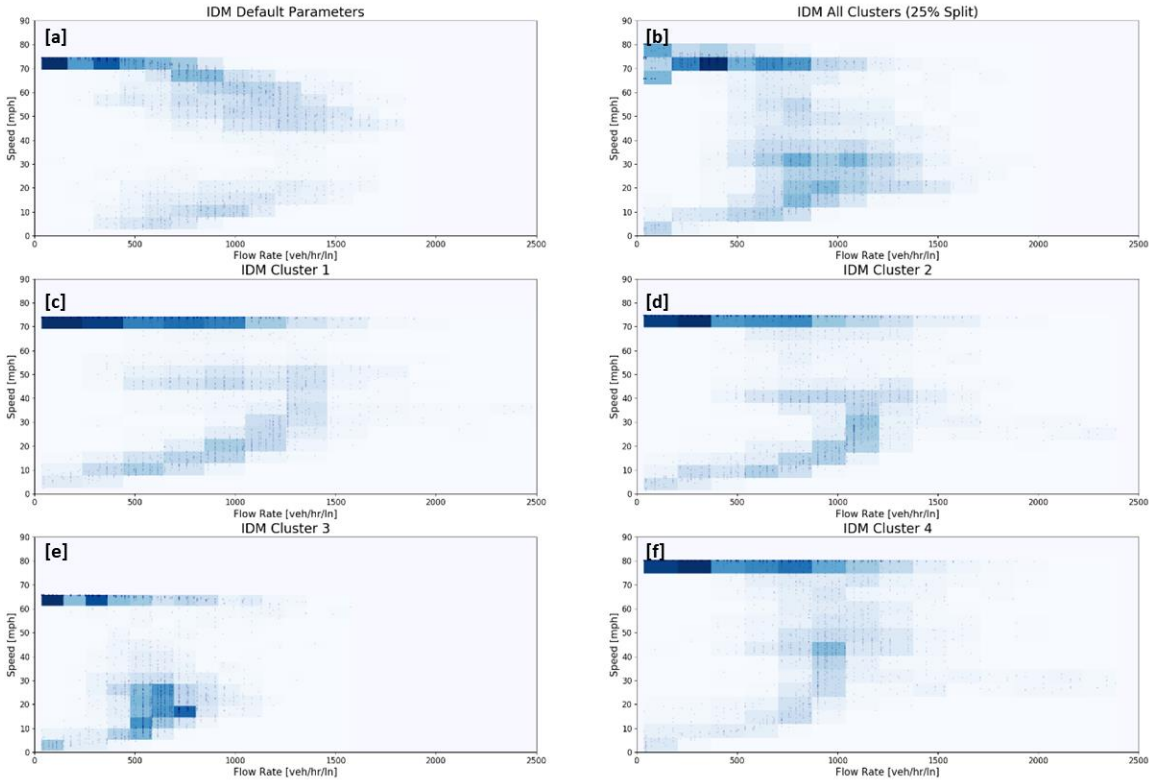


Figure 8.11 Speed-Flow Fundamental Diagrams for Different Proportions of the IDM Homogeneous Driver Groups

The speed-density fundamental diagrams produced using the IDM car-following model are shown in Figure 8.12. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. According to Figure 8.12, the largest jam density was observed when using the default parameter values to simulate the flow of traffic. The smallest jam density was observed when using Cluster 3's parameter coefficients to simulate the flow of traffic; as a reminder, Cluster 3 represents widow(er) drivers.

The relationship between speed and density appears to be mostly linear, as evident by the relatively high goodness-of-fit metrics; this is consistent with Greenshield's theory of a linear relationship between speed and density.

The fundamental diagram produced using Cluster 3's parameter coefficients was the only one to not have speed observations at or above 70 mph; this is consistent with the significantly lower estimate for desired velocity for widow(er) drivers.

The linear regression model constructed using the traffic flow data collected by simulating Cluster 1's representative parameter set, has the steepest slope (i.e., steepest decline from free-flow speed to jam density). This indicates that traffic flow deteriorates more quickly with the parameter set representing single and divorced drivers. The default parameters produced the most gradual slope; this is consistent with what was observed for the Gipps default parameters. Moreover, the goodness-of-fit metric for the linear regression trendline was highest for the default parameters; this suggests the strongest agreement with Greenshield's theory and is the opposite of what was observed for the Gipps model results (see Figure 8.9).

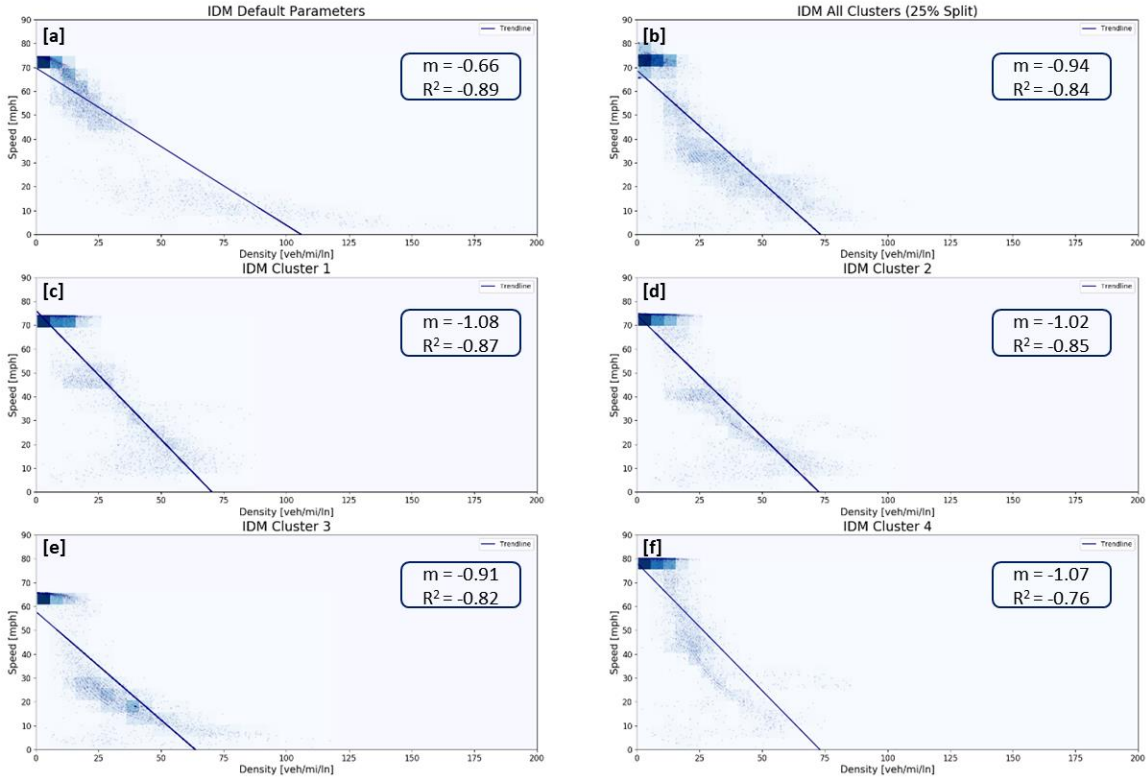


Figure 8.12 Speed-Density Fundamental Diagrams for Different Proportions of the IDM Homogeneous Driver Groups

Finally, Table 8.8 shows the jam densities and capacities of the network (see Figure 8.2) simulated using the IDM car-following behavior according to the different representative parameter sets for the four homogeneous driver groups. The traffic flow simulated with default parameter coefficients produced a significantly higher jam density than the calibrated values. However, the jam densities produced using the calibrated parameters were remarkably similar.

Cluster 3's representative parameter set simulated traffic with the smallest capacity; Cluster 3 represented widow(er) drivers. The remaining calibrated parameters produced higher simulated capacities than what was observed for the default parameters. The highest capacity was observed using Cluster 1's representative parameter set, which represented

single and divorced drivers. Despite similarities in jam density, there is significant variation in capacities as a function of calibration parameter sets used across the five simulations conducted (i.e., a 810 veh/hr/ln difference), which is consistent with what was observed for the Gipps model and underscores the importance of proper calibration (i.e., these parameters significantly impact the outputs of these models).

Table 8.8 Variation in Jam Density and Capacity Across Model Parameters (IDM)

	Jam Density (veh/h)	**Change from Default (%)**	**Capacity (veh/mi)**	**Change from Default (%)**
IDM Default Parameters	105	--	1845	--
IDM All Clusters (25% Split)	73	-30%	1980	7%
IDM Cluster 1	70	-33%	2385	29%
IDM Cluster 2	72	-31%	2285	24%
IDM Cluster 3	63	-40%	1575	-15%
IDM Cluster 4	73	-30%	2285	24%

Wiedemann 99 Car-Following Model Results

The flow-density fundamental diagrams produced using the W99 car-following model are shown in Figure 8.13. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. For all fundamental diagrams produced with the W99 car-following model, there is a clear linear relationship between volume and density in the uncongested regime. There are not strong relationships of density and volume in the congested regime. However, there is more dispersion of data points in the congested flow regime using the default parameters (see Figure 8.13a). This suggests

that there is less uniformity in simulated traffic flow compared to traffic simulated with calibrated parameters. The capacity of the fundamental diagram (see Figure 8.13a) simulated with default parameters also appears to be considerably higher than when calibrated parameters are used to simulate traffic flow (see Figure 8.13b through Figure 8.13f).

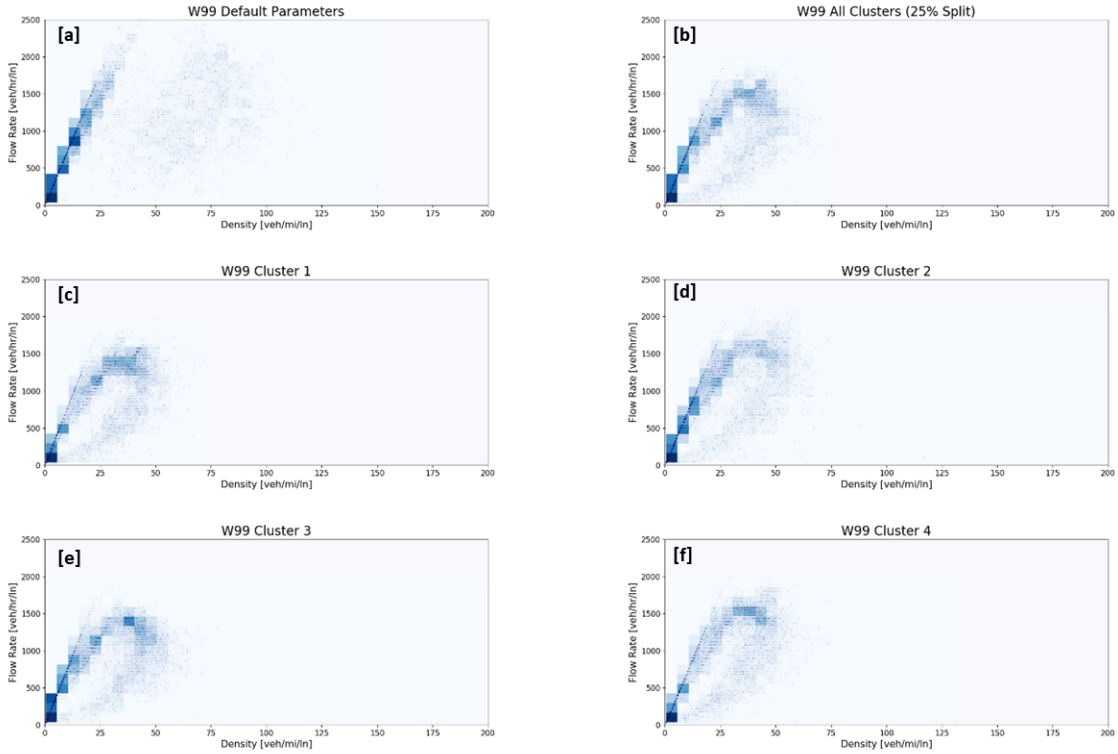


Figure 8.13 Flow-Density Fundamental Diagrams for Different Proportions of the W99 Homogeneous Driver Groups

The speed-flow fundamental diagrams produced using the W99 car-following model are shown in Figure 8.14. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. According to Figure 8.14, when using default parameters, drivers can maintain their desired speeds for longer; this is

evident by darker squares at higher speeds and volumes. Behaviorally speaking, this suggests that drivers simulated with default parameters are less sensitive to surrounding drivers because of the interaction of the parameter coefficients.

More clear patterns in the deterioration of traffic flow are observed when using calibrated model parameters than when using default parameters; this is consistent with observations from Figure 8.13. Additionally, the flow breakdown from uncongested to congested flow is also more easily observable than when default parameters are used (i.e., no significant gaps in the data at moderate speeds).

Patterns from the Cluster 1, Cluster 2, Cluster 3, and Cluster 4 fundamental diagrams (Figure 8.14c through Figure 8.14f) are all represented in the fundamental diagram constructed using equal proportions of clusters (Figure 8.14b). This underscores the value of using multiple parameter sets to simulate diverse driving behavior in naturalistic data.

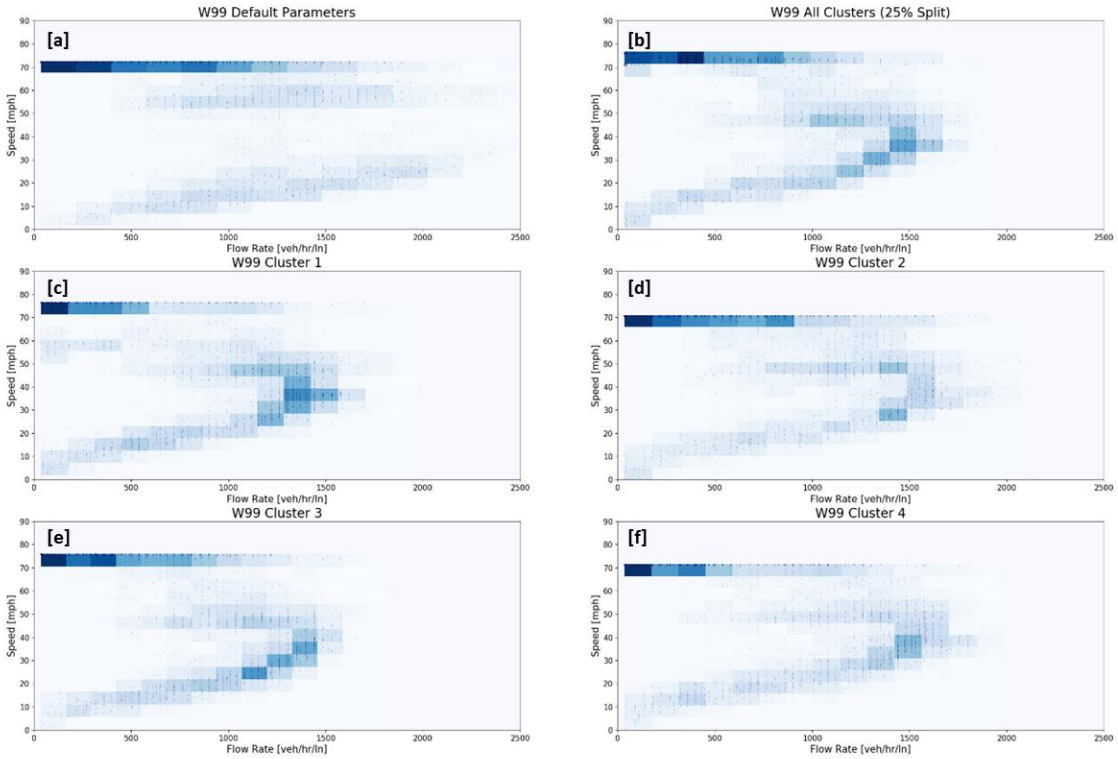


Figure 8.14 Speed-Flow Fundamental Diagrams for Different Proportions of the W99 Homogeneous Driver Groups

The speed-density fundamental diagrams produced using the W99 car-following model to simulate traffic are shown in Figure 8.15. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. As shown in Figure 8.15a, the default parameters produced the largest jam density. The smallest jam density was obtained using Cluster 1's and Cluster 3's calibrated parameter coefficients. The calibrated jam density values were remarkably similar, excluding Cluster 2's results (see Figure 8.15 and Table 8.9).

The linear regression model with the steepest slope, which suggests a quicker traffic deterioration into unstable flow, was observed for Cluster 3's representative parameter

coefficients (see Figure 8.15e). As with the IDM (see Figure 8.9) and Gipps (see Figure 8.12) fundamental diagrams, the default calibration parameters produced the most gradual slope.

The default parameters produced the highest goodness-of-fit metric, suggesting more consistency with Greenshield's theory of speed-density relationships (see Figure 8.15a). This is consistent with what was observed with the IDM model results in Figure 8.12. The smallest goodness-of-fit metric was observed for Cluster 2 (see Figure 8.15d).

Finally, there is a higher concentration of data points simulated using the calibrated parameters compared to the traffic flow simulated with default parameters. This suggests more uniform flow of traffic and is consistent with observations about Figure 8.13.

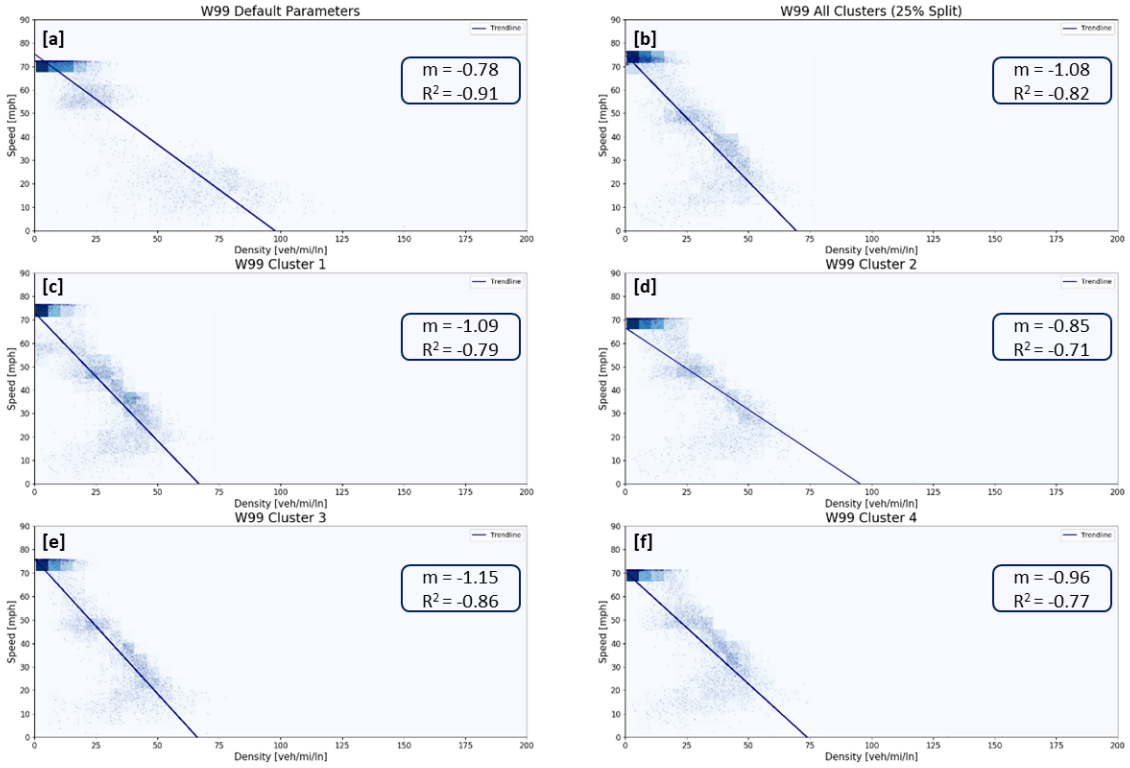


Figure 8.15 Speed-Density Fundamental Diagrams for Different Proportions of the W99 Homogeneous Driver Groups

Table 8.9 shows the jam densities and capacities of the network (see Figure 8.2) simulated using the W99 car-following behavior according to the different representative parameter sets for the four homogeneous driver groups. The jam density is significantly lower when using calibrated parameters compared to default parameters. There is a higher degree of variation of jam densities when using the W99 model to simulate traffic compared to what was observed for the Gipps and IDM models.

The smallest jam density was observed using Cluster 1's and Cluster 3's parameter coefficients to simulate traffic. The largest jam density, obtained using calibrated parameter coefficients, was observed using Cluster 2's parameters. The capacity is significantly lower when using calibrated parameters compared to default parameters. The smallest capacity

was observed when using Cluster 3's calibrated parameters to simulate traffic. The largest capacity was obtained when using Cluster 2's calibrated parameters to simulate traffic. It should be noted that there is significantly less variation in capacities compared to the Gipps and IDM results.

Table 8.9 Variation in Jam Density and Capacity Across Model Parameters (W99)

	Jam Density (veh/h)	Change from Default (%)	Capacity (veh/mi)	Change from Default (%)
W99 Default Parameters	97	--	2495	--
W99 All Clusters (25% Split)	69	-29%	1944	-22%
W99 Cluster 1	66	-32%	1980	-21%
W99 Cluster 2	95	-2%	2070	-17%
W99 Cluster 3	66	-32%	1845	-26%
W99 Cluster 4	73	-25%	1980	-21%

8.4. CONCLUSIONS AND FUTURE RESEARCH

In this chapter, a new framework for calibrating car-following models was explored. This framework takes advantage of existing large-scale trajectory-level datasets, such as the SHRP2 NDS, to shift the calibration burden toward the back-end of the effort, significantly increasing the practicality of accounting for inter-driver differences in microsimulation models. The proposed framework for using census-level data in the calibration process is shown in Figure 8.1.

The proposed framework builds off concepts developed as part of this dissertation—leveraging the homogeneous driver groups, developed in Section 5.3, and the calibration methods for collections of trajectories, developed in Chapter 7.

In practice, the next step would be to define the proportions of driver groups expected to traverse the study area. This step is more art than science, as it is highly improbable to identify the exact drivers traversing a roadway without connected vehicle technology. Thus, given that '*the art of the deal is the compromise*' this dissertation proposes that, in practice, a modeler could use the most recent census of the study area to identify the most probable percentages of different homogeneous driver groups expected to traverse an area.

The challenge with Chapter 8 is validating the accuracy of the proposed framework, as the network used for simulation is fictitious and no ground truth validation data is available. Thus, to illustrate the benefit of the proposed framework, this dissertation simulated traffic flow in VISSIM and showed that different proportions of driver groups produce reasonable microsimulation performance metrics (e.g., capacity, jam density). Data was collected and aggregated at small intervals and forecasted to hourly flow rates to construct fundamental diagrams describing the network.

Through this case study, several notable observations were made:

- For the Gipps and the IDM car-following models, there is significant variation in capacity results as a function of calibration parameter sets used across the five simulations conducted, underscoring the importance of proper calibration (i.e., these parameters significantly impact the outputs of these models). For the W99 model, the capacities were considerably more similar, but the jam densities had increased variation.
- For the simulations involving 100% of a homogeneous driver group, there were often more observable trends in congested conditions than when default parameters were used. Moreover, results obtained simulating the heterogeneous model, which include representation from all four

homogeneous driver groups, contained the trends from each of the homogeneous model runs. This suggests that the heterogeneous model is better capable of capturing the diverse driving behavior in the NDS. This truly underscores the value of this research effort.

- Greenshield's model, or a linear relationship between speed and density, is known to be an oversimplification of traffic flow theory. The simulated traffic flow using the W99 and IDM default parameters produced the highest goodness-of-fit metrics with the linear regression trendline. This may suggest that the default parameters are not representative of naturalistic driver behavior.
- Finally, it was observed that the Gipps, IDM, and W99 default parameters produced the most gradual transition from free-flow speed to jam density, according to the slope of the line of best fit obtained for the speed-density fundamental diagrams. This may suggest that the default parameters were selected for stability and smoothness of traffic flow (i.e., idealized flow) rather than based on real data.

Although ground truth verification of the framework is not possible at this stage of the research, there is considerable evidence provided in this chapter that changing the proportions of homogeneous driver groups represented in the simulation does impact the key performance metrics of interest as outputs of microsimulation models. Thus, this chapter supports further exploration and verification of the calibration framework in future studies, which are planned but outside of the scope of the current dissertation.

Chapter 9: New Method to Capture Diverse Driving Behavior Heterogeneity – An Ensemble Car-Following Model (Task 5)

9.1. MOTIVATION

As discussed in the literature review, it is well documented that individual preferences in driving behavior (e.g., speed, reaction time) are highly variable between and within drivers. Both intra-driver heterogeneity—defined as the same driver changing their behavior as a function of the changing driving environment—and inter-driver heterogeneity—or different drivers behaving differently despite similarities in the driving environment—have been observed in empirical data. Examples of intra-driver heterogeneity include a driver changing their behavior between clear and adverse weather conditions (Hammit et al., 2019) or through the different elements of a work zone (A. L. Berthaume et al., 2018). Examples of inter-driver heterogeneity include different drivers exhibiting different following behaviors on the same stretch of roadway, such as documented differences in the headway selected by passenger car and heavy vehicle drivers (Durrani et al., 2016).

In Chapter 6 of this dissertation, it was shown that different models can capture some drivers' behaviors better than others. That is, there is no one best car-following model to predict the behavior of all drivers and for all scenarios. In the field of data science, selecting the best model to answer a question is a well-known problem. To overcome the limitation of classical methods there exists a suite of methods—called ensemble methods—that use elements from multiple classical methods to ultimately create a more predictive method. This motivates the question explored in this chapter: *can microsimulation model accuracy be improved by using multiple car-following models to capture the diverse driving behavior evident in naturalistic data*? This dissertation hypothesizes that the combination of multiple car-following models will be able to better capture the inter-driver

heterogeneity within the data compared against the original car-following models, used independently.

The remainder of this chapter is organized as follows. Section 9.2 briefly describes the methodology applied to develop the ensemble car-following model. Section 9.3 highlights important results. Finally, conclusions, limitations, and future research are discussed in Section 9.4.

It should be noted that this chapter is not able to determine the accuracy of using this framework for car-following model calibration, compared against traditional methods of car-following model calibration, as the simulation network is hypothetical and ground-truth verification data is not available. However, this chapter does show that different proportions of driver groups do produce considerably different key performance metrics of a microsimulation model (e.g., capacity), which could impact the results of an alternatives assessment and ultimately influence which projects are selected for funding. At a minimum, this chapter successfully provides evidence suggesting that the influence of inter-driver differences on key transportation performance metrics are worthy of additional analysis.

9.2. METHODOLOGY

This chapter describes the creation of the ensemble car-following model. This includes the selection of car-following models, the identification of representative parameter sets, and the development of the microsimulation case study.

Obtaining Trip-Specific Calibration Coefficients

This methodology requires trip-specific calibration coefficients as inputs. In order to obtain these coefficients, high-resolution trajectory data must be available. This chapter makes use of a sample of the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study dataset obtained by the Wyoming Department of Transportation; for additional details on this dataset, please see Section 3.1.3. Next, the trajectory data must be processed to identify constrained driving, or car-following, states. For details on the procedure used as part of this dissertation, see Section 3.2.2. Finally, optimal car-following model calibration coefficients are identified for each trip by solving a mathematical model of the car-following process with a heuristic; for details on the defined nonlinear optimization problem and the genetic algorithm applied as part of this dissertation, see Section 3.2.4. It should be noted that for this dissertation the Gipps, Intelligent Driver Model (IDM), and Wiedemann 99 (W99) car-following models were selected for analysis; the model specifications are detailed in Section 3.2.3.

Segmenting the Data

In Chapter 6, it was shown that some models can capture the driving behavior of certain drivers better than others. This chapter will apply the attribute driver mileage last year, collected as part of the SHRP2 NDS, dataset to segment the data. The car-following model that achieved the lowest median score for a subcategory of the driver mileage last year attribute is used to construct the ensemble car-following model. For ease of reference, the breakdown of data subcategories and selected model is summarized below:

- 0–5k | Modeled using W99;
- 6–9k | Modeled using IDM;
- 10–12k | Modeled using W99;

- 13–15k | Modeled using IDM;
- 16–19k | Modeled using Gipps;
- 20–23k | Modeled using W99; and
- 25k+ | Modeled using Gipps.

Obtaining Representative Car-Following Model Parameter Set

In Chapter 7, different methods were applied to identify the best procedure for obtaining a representative parameter set from a collection of trajectories (i.e., the collection of all trajectories that belong to all drivers that belong in a subcategory of reported miles driven last year). For additional details on this study, the reader is referred to Chapter 7. From this project, it was identified that taking the mean of each independent parameter works best for the IDM and Gipps models, while taking the median of each independent parameter works best for the W99 model.

The optimal parameters for the Gipps, IDM, and W99 models are shown in Table 9.1, Table 9.2, and Table 9.3, respectively.

Table 9.1 Optimal Parameter Values for the 16–19k and 25k+ Subcategories

Model Parameters		Default	16–19k	25k+
Gipps	Desired Velocity [m/s]	35.0	33.8	34.4
	Desired Acceleration [m/s^2]	2.0	1.3	1.2
	Reaction Time [s]	0.7	0.7	0.5
	Desired Deceleration [m/s^2]	-3.0	-2.8	-3.0
	Predicted Maximum Deceleration of Lead Vehicle [m/s^2]	-3.5	-2.6	-2.9
	Minimum Standstill Distance [m]	1.0	4.6	3.9

Table 9.2 Optimal Parameter Values for the 6–9k and 13–15k Subcategories

Model Parameters		Default	6–9k	13–15k
Intelligent Driver Model	Desired Velocity [m/s]	35.0	30.4	32.6
	Free Acceleration Component	4.0	35.8	54.8
	Desired Time Gap [s]	1.5	1.2	0.8
	Jam Distance [m]	2.0	3.2	3.7
	Desired Acceleration [m/s^2]	1.4	0.6	0.9
	Desired Deceleration [m/s^2]	2.0	2.1	2.2

Table 9.3 Optimal Parameter Values for the 0–5k, 10–12k, and 20–24k Subcategories

Model Parameters		Default	0–5k	10–12k	20–24k
W99	CC0: Standstill Distance [m]	1.5	4.0	4.3	3.7
	CC1: Spacing Time [s]	1.3	0.7	1.0	0.6
	CC2: Following Variation, Max Drift [m]	4.0	12.8	11.8	12.5
	CC3: Threshold for Entering Following [s]	-12	-24.0	-23.9	-23.4
	CC4: Negative Following Threshold [m/s]	-0.25	0.0	-0.1	-0.1
	CC5: Positive Following Threshold [m/s]	0.35	0.95	1.2	0.95
	CC6: Speed Dependency of Oscillation [10^{-4} rad/s]	0.0006	1.8	1.3	2.4
	CC7: Oscillation Acceleration [m/s^2]	0.25	1.0	0.9	0.7
	CC8: Standstill Acceleration [m/s^2]	2.0	1.25	1.4	1.3
	CC9: Acceleration at 80kph [m/s^2]	1.5	0.2	0.1	0.2
	Desired Velocity [m/s]	35.0	31.6	31.8	30.9

Microsimulation Analysis

This case study is very similar to the case study conducted in Chapter 8. This dissertation uses a recommended simple weaving segment from the Sixth Edition of the Highway Capacity Manual (Transportation Research Board, 2010), which includes four-

lanes in one-direction with three on- and off-ramps, as shown in Figure 8.2. As with Chapter 8, VISSIM 9 was used to simulate the driver behavior.

The Gipps and IDM car-following model logic are not included in the VISSIM software. Thus, to develop the ensemble car-following model, the External Driver Model Dynamic Linking Library (DLL) Interface of VISSIM was employed. This allows modelers to apply their own driver behavior models into the VISSIM simulation using C/C++ code. For a detailed overview of the VISSIM External Driver Model interface and an example of the IDM car-following model code, please see Section 8.2.

Data collection points were set up at five locations across the network. Data was aggregated at 20 s intervals and forecasted to hourly flow rates. The volume and speed collected via VISSIM were used to obtain traffic density through the fundamental relationship of traffic flow. This allowed for fundamental diagrams to be developed. For the purpose of this case study, eight sets of simulation runs were completed for each car-following model:

- 100% of vehicles assigned to the 0–5k representative car-following model;
- 100% of vehicles assigned to the 6–9k representative car-following model;
- 100% of vehicles assigned to the 10–12k representative car-following model;
- 100% of vehicles assigned to the 13–15k representative car-following model;
- 100% of vehicles assigned to the 16–19k representative car-following model;
- 100% of vehicles assigned to the 20–24k representative car-following model;

- 100% of vehicles assigned to the 25k+ representative car-following model; and
- The Ensemble car-following model, with equal proportions of the previously listed car-following models.

9.3. RESULTS

The flow-density fundamental diagrams are pictured in Figure 9.1. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. For all fundamental diagrams in Figure 9.1, there is a clear linear relationship between volume and density in the uncongested regime. There are not strong relationships between density and volume in the congested regime. However, there are clear differences between the trends of how traffic flow deteriorates between the three different car-following models: Gipps (Figure 9.1e and Figure 9.1g), IDM (Figure 9.1b and Figure 9.1d), and W99 (Figure 9.1a, Figure 9.1c, and Figure 9.1f).

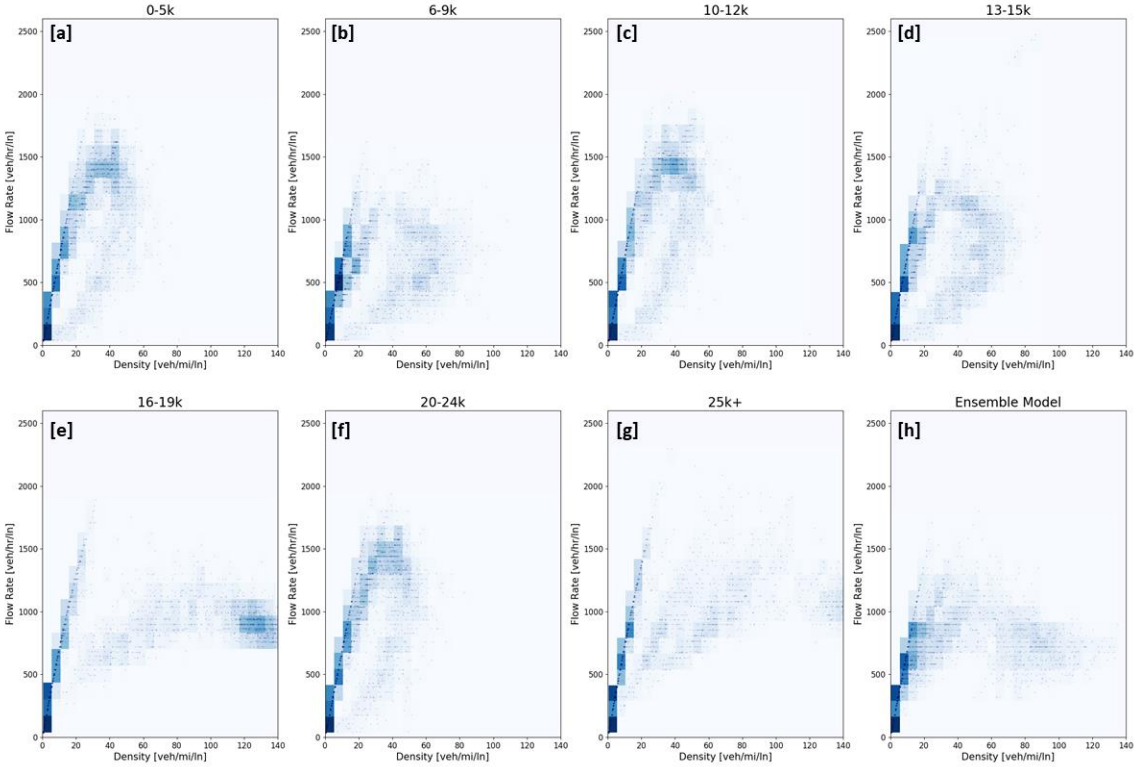


Figure 9.1 Flow-Density Fundamental Diagrams

The speed-flow fundamental diagrams are picture in Figure 9.2. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. The 16–19k (Figure 9.2e) and 25k+ (Figure 9.2g) subcategories have the least clear traffic flow trends in uncongested conditions (i.e., fewer dark squares at high speeds and low volumes). The absence of data at moderate speeds suggests that the traffic flow transitions very quickly from under saturated to oversaturated conditions, as the transitional period is not captured in the data. Moreover, the 16–19k and 25k+ subcategories observed their highest flow rates at relatively higher speeds; this suggest that there is greater instability in queue discharge, which is supported by other observations about the fundamental diagrams.

As shown in Figure 9.2a, Figure 9.2c, and Figure 9.2f, drivers described by the W99 model were able to travel at their desired speeds for longer than when IDM or Gipps were used to simulate traffic flow. This suggests that these drivers are less sensitive to the drivers around them.

Finally, and perhaps most importantly, the Ensemble car-following model (Figure 9.2h) captures traffic flow trends evident in the fundamental diagrams for the individual subcategories (Figure 9.2a through Figure 9.2g). This suggests that the Ensemble model can better capture the diverse driving behaviors compared against any individual car-following model on its own.

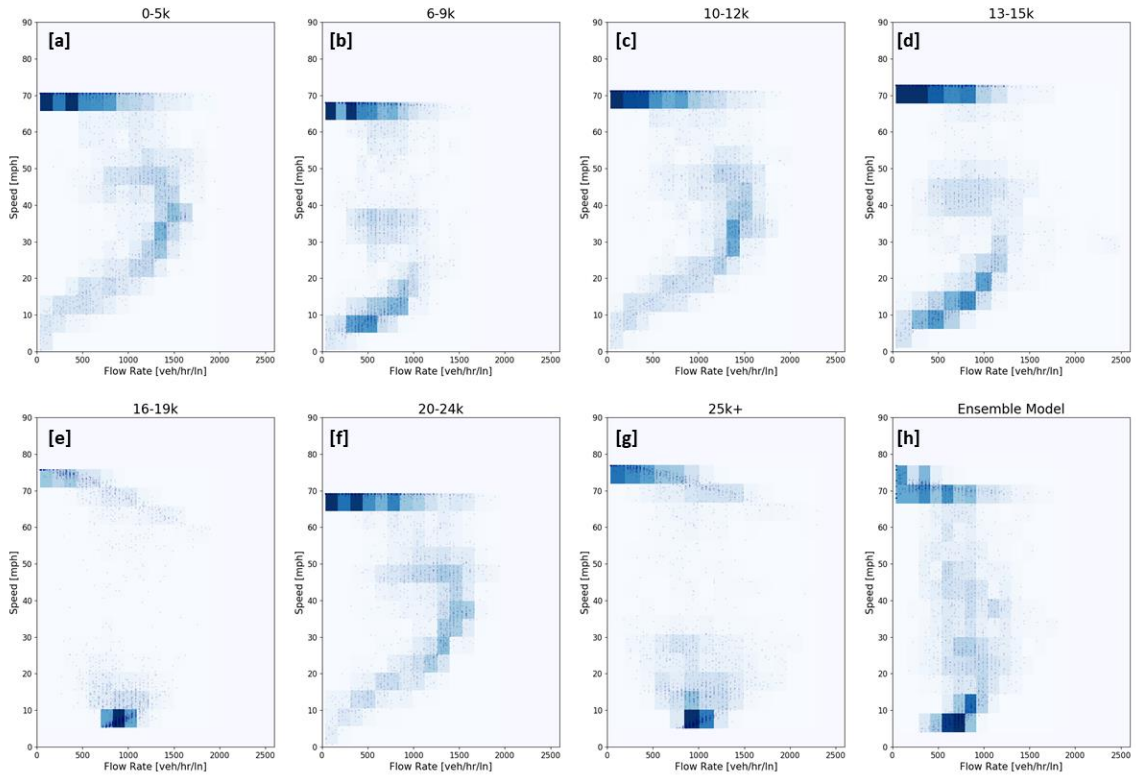


Figure 9.2 Speed-Flow Fundamental Diagrams

The speed-density fundamental diagrams are pictured in Figure 9.3. It should be noted that darker squares represent a high frequency of observed data points; the lighter the squares, the lesser the number of observed data points in that region of the fundamental diagram. The 13–15k subcategory was found to have the steepest slope for the linear regression trendline; this suggests that traffic flow rapidly deteriorates from uncongested to congested conditions.

The 16–19k and 25k+ subcategories were found to have the most gradual slopes, but this is due to the large number of observations at higher densities. Moreover, the 16–19k and 25k+ categories have significantly larger jam densities than the other categories. These categories were simulated using the Gipps model; this is consistent with Chapter 8 observations about the simulated Gipps traffic. This is interpreted as an artifact of the performance of the Gipps car-following model in the VISSIM External Driver Model interface, rather than an observation about drivers that belong in this subcategory of data.

The Ensemble model (see Figure 9.3h) had the highest goodness-of-fit metric. This is likely attributable to the ability of the Ensemble model to produce speed-density points along the spectrum of values from free-flow speed to jam density, unlike the other subcategories, where the data was more concentrated at various points of the line. As observed with Figure 9.2, the Ensemble model captures driving behavior across the complete spectrum of speeds. This suggests that the Ensemble model is better able to capture the diverse driving behavior compared to any individual car-following model.

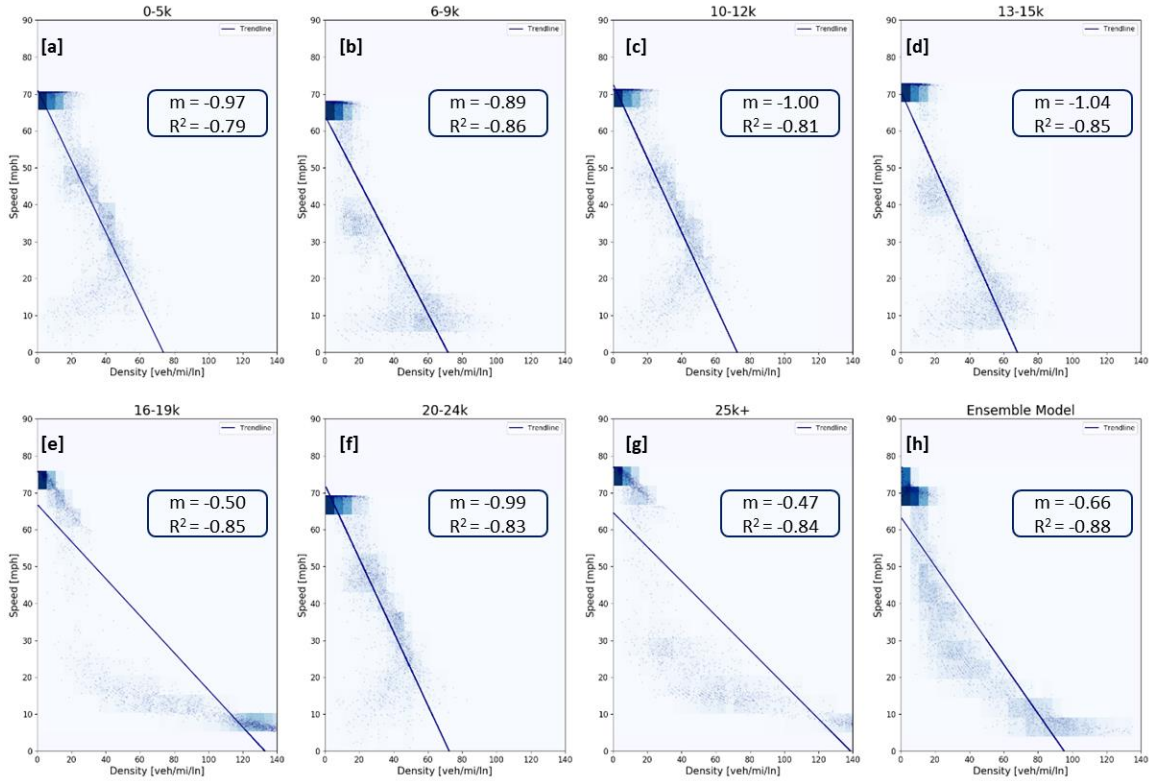


Figure 9.3 Speed-Density Fundamental Diagrams

Finally, the jam densities and capacities of each network are shown in Table 9.4. The jam densities produced by the W99 and IDM car-following models are remarkably close in magnitude. The jam densities produced by the Gipps model are significantly higher than those produced by other car-following models. As discussed in Chapter 8, this is likely because the Gipps model, applied through the External Driver Model interface, is not completely collision free; the Gipps model has an explicit reaction time, where the driver is acting upon information from the time step that is their reaction time in the past. Thus, when the vehicle changes lanes, the driver is not appropriately reacting to their new leading vehicle until a reaction time later and some collisions ensue.

However, the capacities are very different across the different driver groups. Although there is only a five veh/mi/ln difference in jam densities across the different

groups modeled by the W99 and IDM models, there is an 855 veh/hr/ln difference in capacities. Upon inspection of the simulation, these differences seem to be related to how traffic transitions from uncongested to congested conditions. It is a known problem in microsimulation modeling that the car-following model calibration solution may not be unique (i.e., different combinations of parameter values may produce a similar calibration score). Moreover, different sets of calibration parameter coefficients produce similar outputs of microsimulation models (see the results for the 0–5k and 20–24k subcategories of drivers in Table 9.4). However, visual inspection of the simulation and fundamental diagrams confirms that there are differences in how the traffic flow deteriorates attributable to different car-following model functional forms. Although outside of the scope of this dissertation, this concept of exploring the nuanced differences in the deterioration of traffic flow in the congested regime may help to provide insights into uniquely optimal calibration of simulation models.

Table 9.4 Variation in Jam Density and Capacity Across Model Parameters

	Model	**Jam Density (veh/m/ln)**	**Capacity (veh/h/ln)**
0-5k	W99	73	1980
6-9k	IDM	71	1620
10-12k	W99	72	2016
13-15k	IDM	68	2475
16-19k	Gipps	132	1890
20-24k	W99	72	1935
25k	Gipps	138	2295
Ensemble	All	95	1800

9.4. CONCLUSIONS AND FUTURE RESEARCH

This chapter explores the idea of an ensemble car-following model, or a model that makes use of multiple car-following models to best describe the collective behavior of a diverse set of drivers.

Key insights of this chapter include the following:

- As evident in Figure 9.2, the Ensemble car-following model captures traffic flow trends evident in the fundamental diagrams for the individual subcategories. This may suggest that the Ensemble model can better capture the diverse driving behaviors better than any individual car-following model, independently.
- As evident in Figure 9.3, the Ensemble model captures driving behavior across the complete spectrum of speeds. This suggests that the Ensemble model is better able to capture the diverse driving behavior compared to any individual car-following model.
- Although some of the jam densities were very similar, the models and estimated parameter values produced very different fundamental diagrams and capacities. This suggests that the deterioration of traffic flow is different with different car-following models. Although outside of the scope of this dissertation, the study of the deterioration of traffic flow into the congested regime may help to provide insights into uniquely optimal calibration of simulation models.

Unfortunately, this chapter was not able to validate the ensemble model, as there is no ground truth data available about the fictitious network used for simulation. Instead, this chapter was intended to show that the different models and optimal parameter coefficients produce different key performance metrics of microsimulation models, which may

ultimately have an impact on the assessment of transportation alternatives and project selection.

Future research will explore the concept of the Ensemble car-following model on a real-world network, where ground truth verification data can be collected and the validity of the model can be assessed

Chapter 10: Final Conclusions and Recommendations

In August 2018, the author attended and presented some preliminary results of this dissertation at the Seventh International Symposium on Naturalistic Driving Research, in Blacksburg, VA. This two-day conference was dominated by research completed using the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) dataset. The research questions and results presented were predominately safety-oriented. However, as one conference attendee pointedly stated, the majority of the SHRP2 NDS dataset is not made up of crashes or safety critical events: it is boring, baseline driving data that is not of significant value to safety researchers.

However, this *boring, baseline data* is exactly what is needed in the operations discipline. Further analysis of the SHRP2 NDS holds opportunities to explore additional sources of both inter- and intra-driver heterogeneity. Possible future research questions include, but are not limited to:

- Intra- and inter-driver differences attributable to varying levels of congestion (e.g., level-of-service (LOS) A, LOS C, LOS F).
- Intra- and inter-driver differences attributable to facility type (e.g., arterials, freeways).
- Intra- and inter-driver differences attributable to segment categorization (e.g., basic, weaving, merging).
- Intra- and inter-driver differences attributable to operational conditions (e.g., narrow lanes, work zones).
- Intra- and inter-driver differences attributable to time of travel (e.g., weekday, weekend day, peak period, off-peak period).

This dissertation has only begun to scratch the surface of the value in the SHRP2 NDS dataset for operations-focused analysis, modeling, and simulation tools development. It explored the inter-driver differences attributable to driver specific attributes. This dissertation has made three primary contributions to the literature.

First, it identifies the presence of homogeneous driver groups evident in naturalistic data; these groups were identified through the clustering of surrogate measures for driving behavior and validated via classification algorithms that were able to correctly assign drivers to their behavioral groups without information about their driving style.

Secondly, this dissertation introduces a new framework for calibrating the behavioral component of microsimulation models by adjusting the proportions of homogeneous driver groups represented in the population; although this framework was not fully validated, this dissertation was able to illustrate that when different proportions are assumed, it impacts the performance metrics of interest and could influence project selection.

Finally, this dissertation develops an ensemble car-following model, which uses three well-documented and accepted car-following models to capture the diverse driving behavior in naturalistic data.

To provide these three contributions to the transportation community, this dissertation was divided into five tasks:

- Task 1: Evaluate the presence of heterogeneity in the SHRP2 NDS dataset and determine if existing car-following models in the literature are suitable for capturing this heterogeneity at a trip-level.
- Task 2: Characterize the inter-driver heterogeneity evident in the SHRP2 NDS as a function of driver specific attributes (e.g., age, gender).

- Task 3: Identify method(s) to capture the collective behavior of homogeneous driver groups in microsimulation models.
- Task 4: Evaluate the utility of accounting for driver attributes in microsimulation.
- Task 5: Evaluate a new method for capturing inter-driver heterogeneity: an "empirical" driver model.

Through the completion of these five tasks, this dissertation explored eight research questions. The next section provides parting thoughts on these questions.

10.1. Reflections on Research Questions

As defined in Chapter 1, this dissertation aimed to provide insight into eight specific research questions. Final thoughts on the questions and the insights derived through this dissertation research are provided herein.

Research Question 1: Is there evidence of driving behavior heterogeneity in the SHRP2 NDS time-series data?

In Chapter 4, this dissertation analyzed the distributions of trip statistics and calibration parameter values across the 665-trip sample of the SHRP2 NDS used in this dissertation. This Chapter illustrated that there is significant variation across estimated calibration parameter coefficient distributions, even when controlling for facility type and weather conditions. The trip summary statistics are mostly normally distributed. The model calibration parameter coefficients are not normally distributed, aside from the desired velocity parameter. This motivated the use of Kruskal-Wallis one-way analysis of variance (ANOVA) tests to assess the statistical differences in parameter values and model calibration scores.

Research Question 2: Can existing car-following models appropriately capture the diverse driving behavior of different drivers in the sample of SHRP2 NDS?

Yes. Chapter 3 details the joint effort to develop a protocol for processing the SHRP2 NDS to identify the optimal parameter coefficients for three different car-following models in the literature. The calibration procedure identified the set of car-following model parameter coefficients that minimized the root mean square error (RMSE) between the predicted and observed following distance between two vehicles across a trajectory using a genetic algorithm. The genetic algorithm used to solve the nonlinear optimization problem produced reasonable RMSE results, suggesting sufficient model fit. The three car-following models used in this dissertation are substantially different in intuition and functional form. However, there are five model calibration parameters that have similar physical interpretations. Aside from the minimum inter-vehicle spacing parameter, which potentially suffers from data incompleteness, these parameters are remarkably similar in measures of central tendency (e.g., mean, median, mode) and distribution shape (e.g., skewness, kurtosis). This provides support in the validity and accuracy of the calibration procedure, which was optimized using a heuristic instead of an exact solution method.

Research Question 3: Which model(s) best capture the diverse driving behavior recorded in trajectory-level data?

As my adviser would say, it depends! In Chapter 6, Kruskal-Wallis ANOVA tests were applied to assess the differences in calibration score across difference subgroups of drivers, segmented by driver attributes. Different models were able to better capture the driving behavior of certain groups of drivers better than others. As shown in Chapter 6, there is no one car-following model that is the best performing for all drivers and all scenarios. This further supports previously documented observations that in order to

adequately capture the heterogeneity in naturalistic driving data, different sets of optimal calibration parameters and different car-following models may be required. The variation in calibration score was statistically significant across all driver attribute subcategories, except those that comprise the gender attribute. This provides support that different subcategories of driver attributes should be modeled differently.

Research Question 4: Do different hypothesized groups of drivers exhibit statistically significant differences in driving behavior?

Yes. In Section 5.1, Kruskal-Wallis ANOVA tests were conducted to assess the statistical differences in model parameter magnitude across different subcategories of drivers segmented by driver attributes. Section 5.1 provides significant evidence that there are statistically significant differences in driver behavior (i.e., calibrated car-following model parameter estimates) between different subcategories (e.g., male, female) of driver attributes (e.g., gender). Segmenting the data across some attributes, such as driver age and marital status, produced more conclusive trends than other attributes, such as gender and driver mileage last year.

Research Question 5: Do different subgroups of drivers behave sufficiently similarly to be considered one homogeneous group of drivers (i.e., do homogeneous driver groups exist in trajectory-level data?)

Yes. In Section 5.2, clustering and classification algorithms in Weka were applied to show the existence of homogeneous driver groups in the SHRP2 NDS. The successful application of the *Expectations Maximization* clustering algorithm illustrates that there is evidence in naturalistic data that some drivers behave sufficiently similar to one another (and sufficiently different from drivers belonging to a different group) to be considered a

homogeneous group of drivers. To confirm these homogeneous driver groups can be identified via driver attributes, classification algorithms (i.e., *ZeroR*, *OneR*, *PART* Decision Rule and *J48* Decision Tree algorithms) were utilized. To improve model fit, an attribute selection algorithm, *CfsSubsetEval*, was applied to identify the most predictive attributes about a driver. Age and marital status were the most commonly selected attributes, while gender, race, and educational attainment were the least commonly selected attributes. Although the improvements varied depending on the model parameter used for clustering, results indicate that clusters of homogeneous car-following model parameters are correlated with driver attributes (i.e., age). Put succinctly, some drivers drive sufficiently alike to form a cluster of similar behavior. Moreover, driver specific attributes can be carefully utilized to classify drivers into these homogeneous driver groups.

Research Question 6: What methods should be used to obtain a representative set of calibration coefficients for a group of drivers?

In Chapter 7, this dissertation evaluated eight viable methods for obtaining representative sets of calibration parameters. The methods are grouped into three behavioral categories: (i) average behavior, (ii) most frequently observed behavior, or (iii) randomly sampled behavior; moreover, these methods are designed to evaluate the importance of preserving possible correlations between calibration parameters. A 100-trip sample of the SHRP2 NDS was applied to calibrate four common car-following models. A robust validation strategy from the data science literature, 10-fold cross-validation, was implemented to evaluate the performance of the eight methods.

The research findings show that the method that captured the average behavior while preserving correlations between the calibrated model parameters performed the best across all four models; this illustrates the importance of accounting for the underlying

relationships between model parameters. However, methods that adequately captured the average behavior while relaxing the assumption of underlying parameter correlations performed better than all other methods. In other words, although the more computationally burdensome methods produce optimal results, simply taking the mean or median of the distribution of individual parameter coefficients offers a practical approach for generating a representative parameter set. For all models, these methods demonstrated significantly better performance than the default parameter sets. For the Gipps and IDM calibration coefficients, simply taking the average of each independent calibration parameter for the collection of trajectories proved practical and sufficiently accurate; for the W99 calibration coefficients, taking the median of the calibration parameters for the collection of trajectories was found to work best from a practical perspective.

Research Question 7: Can census-level (i.e., driver demographics) data be used alongside the anticipated proportions of driver subgroups to calibrate the car-following behavior of microsimulation models, effectively moving the calibration process to the back-end?

To assess this research question, homogeneous driver groups were identified in the SHRP2 NDS dataset using the framework developed in Chapter 5.2 of this dissertation. The most practical, but accurate, method for each of the car-following models, as identified in Chapter 7 of this dissertation, was applied to identify the representative parameter sets for the homogeneous driver groups. The External Driver Model Dynamic Linking Library (DLL) Interface of VISSIM was used to simulate different assumed proportions of the homogeneous driver groups in the driving population. Key performance metrics, such as flow and speed, were collected to assess differences in network performance as a function of the assumed proportions of homogeneous driver groups.

This case study found that when the Gipps and the IDM car-following models were applied, there is significant variation in capacity results as a function of calibration parameter sets used across the five simulations, underscoring the importance of proper calibration (i.e., these parameters significantly impact the outputs of these models). For the W99 model, the capacities were considerably more similar, but the jam densities had increased variation.

For the simulations involving 100% of a homogeneous driver group, there were often more observable trends in congested conditions than when default parameters were used. Moreover, results obtained simulating the heterogeneous model, which include representation from all four homogeneous driver groups, contained the trends from each of the homogeneous model runs. This suggests that the heterogeneous model is better capable of capturing the diverse driving behavior in the NDS. This truly underscores the value of this research effort.

Finally, it was observed that the Gipps, IDM, and W99 default parameters produced the most gradual transition from free-flow speed to jam density, according to the slope of the line of best fit obtained for the speed-density fundamental diagrams. This may suggest that the default parameters were selected for stability and smoothness of traffic flow (i.e., idealized flow) rather than based on real data. Although ground truth verification of the framework is not possible at this stage of the research, there is considerable evidence provided in this dissertation that changing the proportion of homogeneous driver groups represented in the simulation does impact the key performance metrics of interest as outputs of microsimulation models. Thus, this supports further exploration and verification of the calibration frameworks in future studies, which are planned but outside of the scope of this dissertation.

Research Question 8: Does the diverse driving behavior observed in trajectory-level data require the application of multiple car-following models to more adequately capture the apparent heterogeneity in driving styles?

To assess this research question, the data was segmented by the attribute miles driven last year to create subgroups of drivers. The best method for each of the car-following models, as identified in Chapter 7 of this dissertation, was applied to identify the representative parameter sets for the subgroups of drivers separated by their reported attribute miles driven last year. The External Driver Model DLL Interface of VISSIM was used to simulate different assumed proportions of the driver subgroups in the driving population.

As illustrated in Chapter 9, the Ensemble car-following model captures traffic flow trends evident in the fundamental diagrams for the individual subcategories. This may suggest that the Ensemble model can better capture the diverse driving behaviors better than any individual car-following model, independently. Moreover, the Ensemble model captures driving behavior across the complete spectrum of speeds. This suggests that the Ensemble model is better able to capture the diverse driving behavior compared to any individual car-following model.

Although some of the jam densities were very similar, the models and optimal parameter values produced very different fundamental diagrams and capacities. This suggests that the deterioration of traffic flow is different with different car-following models. Although outside of the scope of this dissertation, the study of the deterioration of traffic flow into the congested regime may help to provide insights into uniquely optimal calibration of simulation models.

Unfortunately, this chapter was not able to validate the ensemble model, as there is no ground truth data available about the fictitious network used for simulation. Instead, this

chapter was intended to show that the different models and optimal parameter coefficients produce different key performance metrics of microsimulation models, which may ultimately have an impact on the assessment of transportation alternatives and project selection.

10.2. Opportunities for Future Research

The original goal of this dissertation was to develop a calibration procedure that could be used to improve the state-of-practice of microsimulation modeling. Although the calibration procedure developed as part of this dissertation was successful, there is a lot of room to improve and increase the practicality of the procedure. First, it is recommended to more deeply explore the uniqueness of the calibration solution. This dissertation observed that different calibrated parameter sets obtain highly similar calibration scores. Thus, as discussed in Chapter 9, studying the deterioration of traffic flow may provide additional support behind the selection of one set of calibration parameters over another.

Moreover, part of the challenge of scaling up the calibration of car-following models with trajectory-level data for application in practice is the computation time required for the calibration procedure, particularly with models with a higher number of calibration coefficients. Later in this research, a literature review revealed methods to minimize the number of calibrated coefficients through the idea of variance-based sensitivity analysis for the IDM (Punzo et al., 2015). Thus, it is recommended to explore this concept on the W99 car-following model, which was the model that took the longest to calibrate.

Finally, prior to calibration, data should be observed for completeness of all traffic conditions (Treiber & Kesting, 2013b). In hindsight, this should have been conducted for

the 665-trip sample used in this dissertation and would have likely reduced the number of parameters required for calibration (i.e., the minimum standstill distance parameter likely lacked sufficient data in most trips for calibration given the nature of the data).

The second area of recommended future research is validating the census-level data calibration framework and the ensemble car-following model. This dissertation used a fictitious network, as described in the Highway Capacity Manual, where validation data (i.e., ground truth capacity) was not available. In future research, it is recommended that this framework and the ensemble car-following model be applied to a real-world network, where the performance can be validated with ground truth data.

Finally, future research recommends continued exploration of the SHRP2 NDS for operations-related questions. This dissertation made use of a 665-trip sample of the larger 5.5 million trip file SHRP2 NDS dataset. This dataset offers an unprecedented opportunity to explore other sources of heterogeneity (e.g., level of congestion) and other driver behaviors (e.g., lane changing). The SHRP2 NDS should be more heavily used in operations-related research before the data expiration date, where all original data must be disposed of per Institutional Research Board requirements.

10.3. CONCLUDING REMARKS

The aging Eisenhower Interstate System and the further diminishing purchasing power of the Federal Gas Tax has created a significant challenge for transportation professionals. Toward the proper stewardship of public funds, the 2015 Fixing America's Surface Transportation (FAST) Act mandates that new transportation projects using federal funding will employ analysis, modeling, and simulation tools to ensure funded alternatives will have positive implications for the network (e.g., reduced congestion, improved safety).

However, simulation models have significant limitations. One of the largest limitations is the availability of appropriate data for model development and calibration. Until recent history, the data available was macroscopic in nature (e.g., x-minute aggregated traffic counts, speed, and travel time). However, high-resolution data collection efforts, such as the SHRP2 NDS and the Next Generation SIMulation (NGSIM) datasets, are providing an unprecedented opportunity to produce more robust funding recommendations through the proper calibration of the underlying model. This dissertation, if nothing else, provides significant evidence of the necessity to more adequately capture behavioral heterogeneity in microsimulation models.

Appendix A: Numerical Results Supporting Section 5.2

Table A.1: Mean Acceleration Segmented by Driver Attributes

Trip Statistic - Mean(Following Vehicle Acceleration) [m/s^2]							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.07	0.14	**0.018**	0.10	**0.002**
	M	261	0.05	0.10		0.05	
Age	a: 20-24	71	0.01	0.15	**0.000**	-0.01	**0.000**
	b: 25-29	209	0.10	0.14		0.13	
	c: 30-34	61	0.05	0.12		0.08	
	d: 35-39	80	0.07	0.09		0.08	
	e: 40-44	66	0.01	0.08		0.02	
	f: 45-59*	47	0.05	0.12		0.04	
	g: 60-69*	71	0.07	0.09		0.07	
	h: 70+*	56	0.04	0.18		0.04	
Race	Caucasian	453	0.07	0.14	**0.045**	0.08	**0.002**
	Not Caucasian	205	0.05	0.10		0.03	
Education	a: No college degree	191	0.12	0.13	**0.00**	0.15	**0.00**
	b: College degree	338	0.05	0.11		0.06	
	c: Graduate Degree	131	-0.01	0.13		0.01	
Marital Status	a: single	219	0.02	0.14	**0.00**	0.02	**0.000**
	b: unmarried partners	49	0.02	0.03		0.02	
	c: married	282	0.09	0.13		0.12	
	d: divorced	89	0.08	0.08		0.10	
	e: widow(er)	19	0.08	0.21		0.09	
Income	a: Under 39k*	93	0.02	0.18	**0.000**	0.02	**0.000**
	b: 40-49k	103	0.03	0.12		0.02	
	c: 50-69k	170	0.15	0.10		0.17	
	d: 70-99k	97	0.06	0.10		0.06	
	e: 100-149k	92	0.02	0.10		0.03	
	f: 150k+	44	-0.07	0.11		-0.05	
HHSize	a: 1	229	0.04	0.13	**0.000**	0.06	**0.000**
	b: 2	151	0.04	0.13		0.06	
	c: 3	191	0.11	0.13		0.14	
	d: 4 or more	89	0.03	0.11		0.02	
Driver Mileage Last Year	a: 0-5k*	40	0.02	0.11	**0.000**	0.00	**0.000**
	b: 6-9k*	41	0.10	0.17		0.09	
	c: 10-12k*	141	0.04	0.13		0.07	
	d: 13-15k*	48	0.00	0.12		0.03	
	e: 16-19k*	45	0.03	0.20		0.07	
	f: 20-23k*	104	0.06	0.08		0.07	
	g: 25k+	170	0.12	0.10		0.14	

Table A.2: Minimum Acceleration Segmented by Driver Attributes

Trip Statistic -Min(Following Vehicle Acceleration) - [m/s^2]							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	-0.88	0.32	**0.000**	-0.84	**0.000**
	M	261	-0.75	0.28		-0.70	
Age	a: 20-24	71	-0.99	0.32	**0.000**	-0.91	**0.000**
	b: 25-29	209	-0.90	0.28		-0.84	
	c: 30-34	61	-0.86	0.28		-0.81	
	d: 35-39	80	-0.77	0.27		-0.71	
	e: 40-44	66	-0.74	0.24		-0.68	
	f: 45-59*	47	-0.79	0.38		-0.71	
	g: 60-69*	71	-0.78	0.32		-0.78	
	h: 70+*	56	-0.64	0.37		-0.58	
Race	Caucasian	453	-0.83	0.33	0.821	-0.79	0.494
	Not Caucasian	205	-0.84	0.27		-0.79	
Education	a: No college degree	191	-0.79	0.31	0.071	-0.76	0.053
	b: College degree	338	-0.85	0.31		-0.82	
	c: Graduate Degree	131	-0.85	0.32		-0.77	
Marital Status	a: single	219	-0.94	0.30	**0.000**	-0.89	**0.000**
	b: unmarried partners	49	-0.72	0.21		-0.67	
	c: married	282	-0.81	0.30		-0.77	
	d: divorced	89	-0.76	0.31		-0.72	
	e: widow(er)	19	-0.58	0.35		-0.55	
Income	a: Under 39k*	93	-0.96	0.39	**0.000**	-0.93	**0.000**
	b: 40-49k	103	-0.86	0.27		-0.83	
	c: 50-69k	170	-0.80	0.27		-0.76	
	d: 70-99k	97	-0.85	0.36		-0.85	
	e: 100-149k	92	-0.76	0.27		-0.68	
	f: 150k+	44	-0.89	0.28		-0.86	
HHSize	a: 1	229	-0.84	0.32	0.054	-0.83	**0.011**
	b: 2	151	-0.85	0.36		-0.80	
	c: 3	191	-0.85	0.29		-0.80	
	d: 4 or more	89	-0.75	0.23		-0.69	
Driver Mileage Last Year	a: 0-5k*	40	-0.98	0.32	**0.000**	-0.96	**0.000**
	b: 6-9k*	41	-0.69	0.34		-0.68	
	c: 10-12k*	141	-0.84	0.32		-0.81	
	d: 13-15k*	48	-0.80	0.30		-0.73	
	e: 16-19k*	45	-0.97	0.37		-0.85	
	f: 20-23k*	104	-0.81	0.33		-0.76	
	g: 25k+	170	-0.78	0.26		-0.71	

Table A.3: Maximum Acceleration Segmented by Driver Attributes

Trip Statistic - Max(Following Vehicle Acceleration) - [m/s^2]							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.90	0.27	**0.000**	0.92	**0.000**
	M	261	0.77	0.21		0.74	
Age	a: 20-24	71	0.87	0.30	**0.000**	0.84	**0.000**
	b: 25-29	209	0.97	0.24		0.98	
	c: 30-34	61	0.82	0.21		0.81	
	d: 35-39	80	0.81	0.23		0.75	
	e: 40-44	66	0.72	0.12		0.71	
	f: 45-59*	47	0.81	0.25		0.81	
	g: 60-69*	71	0.85	0.26		0.83	
	h: 70+*	56	0.65	0.24		0.67	
Race	Caucasian	453	0.86	0.27	0.397	0.84	0.310
	Not Caucasian	205	0.84	0.22		0.81	
Education	a: No college degree	191	0.92	0.25	**0.000**	0.95	**0.000**
	b: College degree	338	0.84	0.24		0.82	
	c: Graduate Degree	131	0.77	0.28		0.72	
Marital Status	a: single	219	0.87	0.29	**0.000**	0.83	**0.000**
	b: unmarried partners	49	0.74	0.12		0.72	
	c: married	282	0.88	0.26		0.90	
	d: divorced	89	0.80	0.19		0.77	
	e: widow(er)	19	0.72	0.27		0.72	
Income	a: Under 39k*	93	0.87	0.33	**0.000**	0.82	**0.000**
	b: 40-49k	103	0.84	0.26		0.82	
	c: 50-69k	170	0.97	0.20		0.99	
	d: 70-99k	97	0.84	0.24		0.84	
	e: 100-149k	92	0.75	0.20		0.72	
	f: 150k+	44	0.69	0.28		0.68	
HHSize	a: 1	229	0.82	0.26	**0.000**	0.78	**0.000**
	b: 2	151	0.82	0.28		0.80	
	c: 3	191	0.95	0.21		0.96	
	d: 4 or more	89	0.76	0.21		0.72	
Driver Mileage Last Year	a: 0-5k*	40	0.92	0.29	**0.000**	0.89	**0.000**
	b: 6-9k*	41	0.84	0.31		0.77	
	c: 10-12k*	141	0.80	0.27		0.80	
	d: 13-15k*	48	0.70	0.17		0.70	
	e: 16-19k*	45	0.93	0.37		0.87	
	f: 20-23k*	104	0.80	0.20		0.75	
	g: 25k+	170	0.92	0.21		0.95	

Table A.4: Mean Relative Velocity Segmented by Driver Attributes

Trip Statistic - Mean(Relative Velocity)[m/s]							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.66	0.97	**0.002**	0.69	**0.005**
	M	260	0.93	1.17		0.86	
Age	a: 20-24	71	0.60	0.80	**0.000**	0.41	**0.000**
	b: 25-29	209	1.04	0.68		0.99	
	c: 30-34	61	0.17	0.53		0.12	
	d: 35-39	80	0.87	0.86		0.78	
	e: 40-44	65	1.84	0.88		2.11	
	f: 45-59*	47	0.35	0.92		0.48	
	g: 60-69*	71	0.43	1.18		0.45	
	h: 70+*	56	-0.02	1.75		0.11	
Race	Caucasian	453	0.63	1.05	**0.000**	0.68	**0.000**
	Not Caucasian	204	1.07	1.00		0.95	
Education	a: No college degree	191	0.61	1.17	**0.000**	0.79	**0.000**
	b: College degree	338	0.63	0.85		0.57	
	c: Graduate Degree	130	1.35	1.18		1.47	
Marital Status	a: single	219	0.73	0.79	**0.000**	0.65	**0.000**
	b: unmarried partners	48	2.21	0.58		2.27	
	c: married	282	0.71	1.00		0.72	
	d: divorced	89	0.75	0.85		0.78	
	e: widow(er)	19	-1.51	1.53		-1.47	
Income	a: Under 39k*	93	0.45	1.49	**0.000**	0.65	**0.000**
	b: 40-49k	103	0.63	0.69		0.67	
	c: 50-69k	170	0.83	0.83		0.87	
	d: 70-99k	97	0.59	0.89		0.51	
	e: 100-149k	91	1.41	1.23		1.72	
	f: 150k+	44	0.52	1.23		0.37	
HHSize	a: 1	229	0.58	1.14	**0.000**	0.78	**0.000**
	b: 2	151	0.59	1.04		0.44	
	c: 3	191	0.81	0.77		0.80	
	d: 4 or more	88	1.46	1.14		1.74	
Driver Mileage Last Year	a: 0-5k*	40	0.85	0.93	**0.000**	0.58	**0.000**
	b: 6-9k*	41	-0.46	1.61		0.01	
	c: 10-12k*	141	0.37	0.90		0.25	
	d: 13-15k*	48	0.60	0.91		0.64	
	e: 16-19k*	45	1.14	1.20		1.10	
	f: 20-23k*	104	0.67	0.88		0.73	
	g: 25k+	169	1.25	0.91		1.12	

Table A.5: Mean Relative Following Distance Segmented by Driver Attributes

Trip Statistic - Mean(Following Distance) [m]							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	49.88	16.23	**0.000**	46.60	**0.000**
	M	261	58.65	15.69		57.90	
Age	a: 20-24	71	48.33	13.43	**0.000**	46.79	**0.000**
	b: 25-29	209	44.53	13.13		42.20	
	c: 30-34	61	48.48	10.20		48.06	
	d: 35-39	80	53.52	13.58		53.53	
	e: 40-44	66	56.14	10.48		55.67	
	f: 45-59*	47	67.79	13.04		67.01	
	g: 60-69*	71	58.05	16.18		58.64	
	h: 70+*	56	76.24	18.13		74.42	
Race	Caucasian	453	55.08	17.48	**0.000**	53.58	**0.000**
	Not Caucasian	205	49.34	13.48		48.41	
Education	a: No college degree	191	54.51	18.90	0.450	51.28	0.507
	b: College degree	338	52.61	16.08		50.06	
	c: Graduate Degree	131	53.39	14.03		53.93	
Marital Status	a: single	219	45.77	13.79	**0.000**	42.91	**0.000**
	b: unmarried partners	49	55.93	7.05		57.00	
	c: married	282	54.59	16.30		53.54	
	d: divorced	89	59.72	16.42		57.76	
	e: widow(er)	19	83.77	15.05		83.00	
Income	a: Under 39k*	93	54.00	19.95	**0.000**	49.73	**0.000**
	b: 40-49k	103	46.86	17.40		41.15	
	c: 50-69k	170	49.92	13.82		47.92	
	d: 70-99k	97	56.23	18.42		55.16	
	e: 100-149k	92	56.91	12.21		57.26	
	f: 150k+	44	60.17	16.65		56.79	
HHSize	a: 1	229	53.83	19.03	**0.000**	50.84	**0.000**
	b: 2	151	57.15	17.10		55.37	
	c: 3	191	49.47	13.47		47.76	
	d: 4 or more	89	53.75	13.04		54.28	
Driver Mileage Last Year	a: 0-5k*	40	47.85	13.14	**0.000**	45.89	**0.000**
	b: 6-9k*	41	66.83	24.27		71.03	
	c: 10-12k*	141	56.14	17.38		51.59	
	d: 13-15k*	48	58.86	14.75		58.70	
	e: 16-19k*	45	50.82	16.61		52.95	
	f: 20-23k*	104	57.39	14.86		55.83	
	g: 25k+	170	51.06	13.60		52.07	

Table A.6: Mean Headway Segmented by Driver Attributes

Trip Statistic - Mean(Headway) [s]							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	1.98	0.76	0.160	1.78	**0.004**
	M	261	2.06	0.63		1.89	
Age	a: 20-24	71	2.16	0.67	**0.000**	2.03	**0.000**
	b: 25-29	209	1.65	0.56		1.55	
	c: 30-34	61	1.84	0.47		1.70	
	d: 35-39	80	1.89	0.50		1.78	
	e: 40-44	66	1.78	0.35		1.70	
	f: 45-59*	47	2.59	0.59		2.47	
	g: 60-69*	71	2.34	0.74		2.32	
	h: 70+*	56	2.89	0.85		2.87	
Race	Caucasian	453	2.02	0.73	0.481	1.87	0.325
	Not Caucasian	205	1.98	0.67		1.77	
Education	a: No college degree	191	1.95	0.80	**0.000**	1.77	**0.000**
	b: College degree	338	2.11	0.69		1.99	
	c: Graduate Degree	131	1.84	0.59		1.72	
Marital Status	a: single	219	1.93	0.67	**0.000**	1.78	**0.000**
	b: unmarried partners	49	1.72	0.21		1.70	
	c: married	282	1.96	0.70		1.80	
	d: divorced	89	2.20	0.62		2.10	
	e: widow(er)	19	3.45	0.82		3.16	
Income	a: Under 39k*	93	2.21	0.89	**0.000**	1.98	**0.000**
	b: 40-49k	103	1.92	0.70		1.67	
	c: 50-69k	170	1.81	0.63		1.71	
	d: 70-99k	97	2.27	0.76		2.19	
	e: 110-149k	92	1.92	0.59		1.79	
	f: 150k+	44	2.20	0.72		1.90	
HHSize	a: 1	229	2.14	0.82	**0.000**	1.94	**0.000**
	b: 2	151	2.16	0.70		2.09	
	c: 3	191	1.83	0.62		1.75	
	d: 4 or more	89	1.82	0.46		1.76	
Driver Mileage Last Year	a: 0-5k*	40	1.88	0.62	**0.000**	1.76	**0.000**
	b: 6-9k*	41	2.71	0.98		2.65	
	c: 10-12k*	141	2.15	0.68		2.03	
	d: 13-15k*	48	2.19	0.55		2.13	
	e: 16-19k*	45	1.87	0.74		1.81	
	f: 20-23k*	104	2.19	0.70		2.06	
	g: 25k+	170	1.67	0.47		1.64	

Table A.7: Estimated Wiedemann 99 Standstill Distance (cc0) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Standstill Distance [m] - cc0							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	4.52	2.47	0.118	4.50	0.079
	M	261	4.22	2.47		3.80	
Age	a: 20-24	71	4.52	2.29	**0.005**	4.50	**0.010**
	b: 25-29	209	4.20	2.33		4.10	
	c: 30-34	61	5.43	2.59		5.00	
	d: 35-39	80	3.86	2.21		3.60	
	e: 40-44	66	4.94	2.64		4.75	
	f: 45-59*	47	4.23	2.74		4.00	
	g: 60-69*	71	4.19	2.53		3.80	
	h: 70+*	56	4.43	2.68		4.20	
Race	Caucasian	453	3.89	2.33	**0.000**	3.70	**0.000**
	Not Caucasian	205	5.49	2.43		5.50	
Education	a: No college degree	191	3.87	2.33	**0.001**	3.60	**0.001**
	b: College degree	338	4.72	2.51		4.60	
	c: Graduate Degree	131	4.39	2.47		4.20	
Marital Status	a: single	219	5.11	2.36	**0.000**	5.00	**0.000**
	b: unmarried partners	49	5.17	2.40		4.90	
	c: married	282	3.80	2.39		3.55	
	d: divorced	89	4.08	2.48		3.70	
	e: widow(er)	19	4.91	2.52		5.30	
Income	a: Under 39k*	93	4.11	2.25	**0.000**	4.20	**0.000**
	b: 40-49k	103	5.65	2.29		5.50	
	c: 50-69k	170	3.77	2.49		3.40	
	d: 70-99k	97	4.65	2.40		4.40	
	e: 100-149k	92	4.60	2.53		4.45	
	f: 150k+	44	4.58	2.36		4.25	
HHSize	a: 1	229	4.64	2.44	**0.003**	4.60	**0.002**
	b: 2	151	4.40	2.60		4.10	
	c: 3	191	3.89	2.32		3.80	
	d: 4 or more	89	4.92	2.49		4.80	
Driver Mileage Last Year	a: 0-5k*	40	4.20	2.17	0.802	4.00	0.859
	b: 6-9k*	41	4.16	2.51		4.20	
	c: 10-12k*	141	4.31	2.50		4.30	
	d: 13-15k*	48	4.55	2.83		4.35	
	e: 16-19k*	45	4.56	2.53		4.20	
	f: 20-23k*	104	4.10	2.47		3.65	
	g: 25k+	170	4.04	2.32		3.85	

Table A.8: Estimated Wiedemann 99 Spacing Time (cc1) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Spacing Time [s] - cc1							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.84	0.56	0.456	0.65	0.295
	M	261	0.81	0.55		0.60	
Age	a: 20-24	71	0.88	0.39	**0.000**	0.80	**0.000**
	b: 25-29	209	0.58	0.27		0.50	
	c: 30-34	61	1.06	0.39		1.00	
	d: 35-39	80	0.60	0.31		0.50	
	e: 40-44	66	0.68	0.35		0.60	
	f: 45-59*	47	1.08	0.51		1.00	
	g: 60-69*	71	1.03	0.64		0.90	
	h: 70+*	56	1.44	1.09		1.25	
Race	Caucasian	453	0.84	0.59	0.251	0.60	0.646
	Not Caucasian	205	0.78	0.45		0.60	
Education	a: No college degree	191	0.85	0.74	0.073	0.60	**0.001**
	b: College degree	338	0.85	0.45		0.80	
	c: Graduate Degree	131	0.72	0.47		0.60	
Marital Status	a: single	219	0.78	0.42	**0.000**	0.70	**0.000**
	b: unmarried partners	49	0.63	0.36		0.60	
	c: married	282	0.84	0.53		0.65	
	d: divorced	89	0.72	0.49		0.50	
	e: widow(er)	19	2.16	1.12		1.80	
Income	a: Under 39k*	93	0.99	0.82	**0.000**	0.80	**0.000**
	b: 40-49k	103	0.91	0.62		0.70	
	c: 50-69k	170	0.68	0.33		0.60	
	d: 70-99k	97	0.88	0.53		0.80	
	e: 100-149k	92	0.80	0.50		0.60	
	f: 150k+	44	1.13	0.57		1.00	
HHSize	a: 1	229	0.82	0.69	**0.000**	0.60	**0.000**
	b: 2	151	1.00	0.58		0.90	
	c: 3	191	0.71	0.32		0.60	
	d: 4 or more	89	0.77	0.46		0.60	
Driver Mileage Last Year	a: 0-5k*	40	0.73	0.44	**0.000**	0.70	**0.000**
	b: 6-9k*	41	1.54	1.03		1.30	
	c: 10-12k*	141	1.06	0.51		1.00	
	d: 13-15k*	48	0.99	0.64		0.80	
	e: 16-19k*	45	0.62	0.46		0.50	
	f: 20-23k*	104	0.76	0.49		0.60	
	g: 25k+	170	0.63	0.32		0.60	

Table A.9: Estimated Wiedemann 99 Following Variation, Maximum Drift (cc2) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Following Variation, Max Drift [m] - cc2							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	10.45	4.88	0.932	12.65	0.599
	M	261	10.41	4.85		12.70	
Age	a: 20-24	71	10.85	4.42	**0.036**	13.50	0.179
	b: 25-29	209	10.70	4.92		13.30	
	c: 30-34	61	10.88	4.38		12.80	
	d: 35-39	80	9.80	5.53		12.60	
	e: 40-44	66	11.04	4.46		13.05	
	f: 45-59*	47	8.13	5.33		6.60	
	g: 60-69*	71	10.27	4.92		12.50	
	h: 70+*	56	10.76	4.22		11.75	
Race	Caucasian	453	9.95	5.03	**0.000**	12.00	**0.000**
	Not Caucasian	205	11.60	4.21		13.80	
Education	a: No college degree	191	9.99	5.04	0.061	11.90	0.690
	b: College degree	338	10.38	4.90		12.80	
	c: Graduate Degree	131	11.28	4.35		13.20	
Marital Status	a: single	219	11.17	4.40	**0.002**	13.50	**0.012**
	b: unmarried partners	49	11.46	4.32		13.50	
	c: married	282	10.10	4.99		12.10	
	d: divorced	89	9.00	5.46		10.20	
	e: widow(er)	19	11.22	4.51		13.40	
Income	a: Under 39k*	93	10.67	4.45	**0.003**	12.20	**0.018**
	b: 40-49k	103	11.86	4.00		13.80	
	c: 50-69k	170	9.95	5.26		12.45	
	d: 70-99k	97	9.53	5.06		11.20	
	e: 100-149k	92	10.52	4.79		12.75	
	f: 150k+	44	11.77	3.75		13.25	
HHSize	a: 1	229	10.63	4.85	0.573	13.00	0.463
	b: 2	151	10.28	4.86		12.40	
	c: 3	191	10.15	4.93		12.10	
	d: 4 or more	89	10.90	4.72		13.50	
Driver Mileage Last Year	a: 0-5k*	40	10.88	4.62	0.188	12.75	0.339
	b: 6-9k*	41	11.84	3.90		13.60	
	c: 10-12k*	141	10.21	4.78		11.80	
	d: 13-15k*	48	9.38	4.76		11.10	
	e: 16-19k*	45	10.58	4.53		11.60	
	f: 20-23k*	104	9.60	5.40		12.45	
	g: 25k+	170	9.95	5.21		12.10	

Table A.10: Estimated Wiedemann 99 Threshold for Entering 'Following' (cc3) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Threshold for Entering 'Following' [s] - cc3							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	-22.58	5.18	**0.002**	-25.00	**0.015**
	M	261	-21.21	6.25		-23.60	
Age	a: 20-24	71	-22.95	5.19	**0.000**	-25.70	**0.000**
	b: 25-29	209	-23.15	4.83		-25.30	
	c: 30-34	61	-22.34	4.89		-23.80	
	d: 35-39	80	-20.32	5.97		-21.85	
	e: 40-44	66	-22.54	6.03		-25.60	
	f: 45-59*	47	-19.98	6.96		-21.00	
	g: 60-69*	71	-21.40	5.99		-23.90	
	h: 70+*	56	-20.80	6.33		-23.00	
Race	Caucasian	453	-21.64	5.88	**0.007**	-24.10	**0.022**
	Not Caucasian	205	-22.92	5.09		-25.10	
Education	a: No college degree	191	-22.31	5.28	0.510	-24.60	0.770
	b: College degree	338	-21.79	5.97		-24.30	
	c: Graduate Degree	131	-22.29	5.42		-24.60	
Marital Status	a: single	219	-22.83	5.20	**0.032**	-25.10	**0.046**
	b: unmarried partners	49	-22.21	6.32		-25.60	
	c: married	282	-21.88	5.57		-24.30	
	d: divorced	89	-20.72	6.50		-23.40	
	e: widow(er)	19	-20.59	5.54		-20.20	
Income	a: Under 39k*	93	-21.71	5.55	**0.009**	-23.90	**0.020**
	b: 40-49k	103	-22.95	5.22		-25.00	
	c: 50-69k	170	-23.01	4.77		-25.10	
	d: 70-99k	97	-20.49	6.71		-22.80	
	e: 100-149k	92	-22.01	5.90		-24.75	
	f: 150k+	44	-22.38	5.63		-24.50	
HHSize	a: 1	229	-22.16	5.46	0.273	-24.30	0.457
	b: 2	151	-21.38	6.13		-23.80	
	c: 3	191	-22.55	5.30		-24.90	
	d: 4 or more	89	-21.76	6.13		-24.90	
Driver Mileage Last Year	a: 0-5k*	40	-21.99	5.30	0.179	-23.95	0.122
	b: 6-9k*	41	-21.67	5.57		-24.20	
	c: 10-12k*	141	-21.27	6.18		-23.90	
	d: 13-15k*	48	-21.09	6.25		-23.65	
	e: 16-19k*	45	-23.55	3.93		-25.00	
	f: 20-23k*	104	-21.05	6.06		-23.35	
	g: 25k+	170	-22.28	5.75		-24.95	

Table A.11: Estimated Wiedemann 99 Negative Following Threshold (cc4) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Negative Following Threshold [m/s] - cc4							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	-0.54	1.10	**0.005**	0.00	**0.006**
	M	261	-0.81	1.34		-0.10	
Age	a: 20-24	71	-0.46	1.08	**0.000**	0.00	**0.000**
	b: 25-29	209	-0.44	0.96		0.00	
	c: 30-34	61	-0.57	1.17		0.00	
	d: 35-39	80	-0.77	1.31		-0.10	
	e: 40-44	66	-0.32	0.71		0.00	
	f: 45-59*	47	-1.09	1.55		-0.30	
	g: 60-69*	71	-0.74	1.21		-0.10	
	h: 70+*	56	-1.52	1.66		-0.75	
Race	Caucasian	453	-0.76	1.31	**0.000**	-0.10	**0.000**
	Not Caucasian	205	-0.39	0.84		0.00	
Education	a: No college degree	191	-0.77	1.39	0.074	-0.10	0.133
	b: College degree	338	-0.66	1.18		-0.10	
	c: Graduate Degree	131	-0.46	0.97		0.00	
Marital Status	a: single	219	-0.42	0.95	**0.000**	0.00	**0.000**
	b: unmarried partners	49	-0.25	0.52		0.00	
	c: married	282	-0.77	1.34		-0.10	
	d: divorced	89	-0.80	1.23		-0.20	
	e: widow(er)	19	-1.74	1.84		-1.10	
Income	a: Under 39k*	93	-0.76	1.30	0.406	-0.10	0.088
	b: 40-49k	103	-0.44	1.00		0.00	
	c: 50-69k	170	-0.60	1.25		-0.10	
	d: 70-99k	97	-0.64	1.03		-0.10	
	e: 100-149k	92	-0.66	1.31		0.00	
	f: 150k+	44	-0.82	1.38		0.00	
HHSize	a: 1	229	-0.74	1.24	0.470	-0.10	0.419
	b: 2	151	-0.64	1.19		-0.10	
	c: 3	191	-0.62	1.20		-0.10	
	d: 4 or more	89	-0.51	1.18		0.00	
Driver Mileage Last Year	a: 0-5k*	40	-0.40	0.95	**0.031**	0.00	0.248
	b: 6-9k*	41	-1.03	1.53		-0.20	
	c: 10-12k*	141	-0.87	1.42		-0.10	
	d: 13-15k*	48	-0.84	1.49		-0.10	
	e: 16-19k*	45	-0.56	1.18		0.00	
	f: 20-23k*	104	-0.77	1.21		-0.10	
	g: 25k+	170	-0.49	1.01		-0.10	

Table A.12: Estimated Wiedemann 99 Positive Following Threshold (cc5) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Positive Following Threshold [m/s] - cc5							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	1.31	1.27	0.734	0.80	0.970
	M	261	1.34	1.33		0.90	
Age	a: 20-24	71	1.15	1.15	**0.000**	0.80	**0.001**
	b: 25-29	209	1.03	1.02		0.70	
	c: 30-34	61	1.58	1.36		1.10	
	d: 35-39	80	1.34	1.20		1.00	
	e: 40-44	66	1.09	1.18		0.60	
	f: 45-59*	47	2.00	1.61		1.50	
	g: 60-69*	71	1.55	1.60		0.80	
	h: 70+*	56	1.72	1.46		1.45	
Race	Caucasian	453	1.39	1.32	**0.040**	1.00	0.070
	Not Caucasian	205	1.16	1.21		0.70	
Education	a: No college degree	191	1.40	1.35	0.149	0.90	0.390
	b: College degree	338	1.35	1.33		0.90	
	c: Graduate Degree	131	1.13	1.08		0.80	
Marital Status	a: single	219	1.17	1.17	**0.004**	0.80	**0.015**
	b: unmarried partners	49	0.95	1.09		0.60	
	c: married	282	1.42	1.35		1.00	
	d: divorced	89	1.41	1.37		0.90	
	e: widow(er)	19	2.05	1.47		1.80	
Income	a: Under 39k*	93	1.29	1.31	0.644	0.80	0.486
	b: 40-49k	103	1.40	1.33		0.90	
	c: 50-69k	170	1.20	1.16		0.70	
	d: 70-99k	97	1.41	1.42		0.90	
	e: 100-149k	92	1.37	1.41		0.80	
	f: 150k+	44	1.51	1.32		1.05	
HHSize	a: 1	229	1.21	1.24	**0.006**	0.80	**0.021**
	b: 2	151	1.64	1.45		1.10	
	c: 3	191	1.22	1.20		0.70	
	d: 4 or more	89	1.26	1.25		0.80	
Driver Mileage Last Year	a: 0-5k*	40	1.46	1.40	**0.012**	0.95	**0.040**
	b: 6-9k*	41	1.36	1.26		1.00	
	c: 10-12k*	141	1.62	1.45		1.20	
	d: 13-15k*	48	1.69	1.54		1.05	
	e: 16-19k*	45	1.02	1.05		0.70	
	f: 20-23k*	104	1.37	1.34		0.95	
	g: 25k+	170	1.14	1.15		0.70	

Table A.13: Estimated Wiedemann 99 Speed Dependency of Oscillation (cc6) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Speed Dependency of Oscillation [10^{-4} rad/s] - cc6							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	2.24	2.19	**0.016**	1.60	**0.014**
	M	261	2.69	2.48		2.00	
Age	a: 20-24	71	1.74	1.91	**0.000**	1.10	**0.001**
	b: 25-29	209	2.31	2.01		1.90	
	c: 30-34	61	1.83	2.23		1.10	
	d: 35-39	80	3.32	2.87		2.70	
	e: 40-44	66	2.14	1.55		2.00	
	f: 45-59*	47	2.28	2.11		2.00	
	g: 60-69*	71	2.80	2.75		1.90	
	h: 70+*	56	2.99	2.89		2.00	
Race	Caucasian	453	2.59	2.46	**0.004**	1.90	**0.028**
	Not Caucasian	205	2.03	1.92		1.60	
Education	a: No college degree	191	2.61	2.38	0.389	2.00	0.075
	b: College degree	338	2.34	2.41		1.55	
	c: Graduate Degree	131	2.34	1.94		2.00	
Marital Status	a: single	219	2.02	1.98	**0.000**	1.40	**0.002**
	b: unmarried partners	49	2.17	1.55		2.00	
	c: married	282	2.50	2.39		2.00	
	d: divorced	89	3.35	2.90		2.40	
	e: widow(er)	19	2.24	2.50		1.50	
Income	a: Under 39k*	93	2.29	2.21	0.281	1.50	0.092
	b: 40-49k	103	1.96	2.16		1.30	
	c: 50-69k	170	2.34	2.19		1.90	
	d: 70-99k	97	2.61	2.39		2.00	
	e: 100-149k	92	2.37	1.99		2.10	
	f: 150k+	44	1.88	1.96		1.40	
HHSize	a: 1	229	2.64	2.51	0.121	1.90	0.596
	b: 2	151	2.55	2.63		1.60	
	c: 3	191	2.19	2.05		1.80	
	d: 4 or more	89	2.13	1.64		1.90	
Driver Mileage Last Year	a: 0-5k*	40	2.68	2.61	0.075	1.80	0.131
	b: 6-9k*	41	2.03	2.00		1.40	
	c: 10-12k*	141	2.34	2.53		1.30	
	d: 13-15k*	48	2.56	2.66		1.85	
	e: 16-19k*	45	2.15	1.77		1.70	
	f: 20-23k*	104	3.13	2.89		2.35	
	g: 25k+	170	2.36	1.86		2.10	

Table A.14: Estimated Wiedemann 99 Oscillatory Acceleration (cc7) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Oscillation Acceleration [m/s^2] - cc7							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	1.71	1.89	0.107	1.10	**0.001**
	M	261	1.46	1.91		0.70	
Age	a: 20-24	71	1.56	1.50	**0.004**	1.10	**0.000**
	b: 25-29	209	1.77	1.89		1.20	
	c: 30-34	61	2.11	2.14		1.10	
	d: 35-39	80	1.55	2.04		0.70	
	e: 40-44	66	1.83	1.83		1.10	
	f: 45-59*	47	0.62	0.80		0.30	
	g: 60-69*	71	1.42	1.89		0.70	
	h: 70+*	56	1.43	2.32		0.30	
Race	Caucasian	453	1.46	1.93	**0.002**	0.70	**0.000**
	Not Caucasian	205	1.96	1.78		1.40	
Education	a: No college degree	191	1.61	2.17	0.873	0.60	**0.024**
	b: College degree	338	1.59	1.78		1.10	
	c: Graduate Degree	131	1.69	1.78		1.00	
Marital Status	a: single	219	1.73	1.67	0.332	1.30	**0.000**
	b: unmarried partners	49	1.95	1.83		1.60	
	c: married	282	1.49	1.96		0.70	
	d: divorced	89	1.48	1.95		0.50	
	e: widow(er)	19	1.98	3.00		0.30	
Income	a: Under 39k*	93	1.31	1.64	**0.019**	0.80	**0.001**
	b: 40-49k	103	2.13	1.97		1.40	
	c: 50-69k	170	1.69	2.04		0.80	
	d: 70-99k	97	1.30	1.65		0.90	
	e: 100-149k	92	1.67	1.87		1.00	
	f: 150k+	44	1.56	1.77		1.00	
HHSize	a: 1	229	1.56	1.86	0.894	1.10	0.509
	b: 2	151	1.58	1.96		0.80	
	c: 3	191	1.66	1.96		0.90	
	d: 4 or more	89	1.71	1.77		1.10	
Driver Mileage Last Year	a: 0-5k*	40	1.66	1.96	0.251	1.00	0.685
	b: 6-9k*	41	1.99	2.43		0.80	
	c: 10-12k*	141	1.49	1.83		0.90	
	d: 13-15k*	48	1.02	1.42		0.70	
	e: 16-19k*	45	1.39	1.70		0.70	
	f: 20-23k*	104	1.50	1.94		0.70	
	g: 25k+	170	1.73	2.05		0.80	

Table A.15: Estimated Wiedemann 99 Standstill Acceleration (cc8) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Standstill Acceleration [m/s^2] - cc8							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	2.09	1.88	0.246	1.40	0.180
	M	261	2.26	1.91		1.50	
Age	a: 20-24	71	1.82	1.65	0.084	1.40	0.058
	b: 25-29	209	2.27	2.04		1.60	
	c: 30-34	61	2.19	1.85		1.40	
	d: 35-39	80	2.22	1.80		1.85	
	e: 40-44	66	2.41	1.96		1.85	
	f: 45-59*	47	2.64	1.99		2.40	
	g: 60-69*	71	1.87	1.82		1.10	
	h: 70+*	56	1.68	1.67		0.95	
Race	Caucasian	453	2.23	1.91	0.130	1.60	0.105
	Not Caucasian	205	1.99	1.88		1.30	
Education	a: No college degree	191	2.22	1.95	0.112	1.60	0.130
	b: College degree	338	2.02	1.82		1.35	
	c: Graduate Degree	131	2.41	1.99		1.80	
Marital Status	a: single	219	2.05	1.88	0.163	1.40	0.151
	b: unmarried partners	49	2.27	1.85		1.80	
	c: married	282	2.33	1.98		1.70	
	d: divorced	89	1.95	1.78		1.30	
	e: widow(er)	19	1.51	1.39		0.70	
Income	a: Under 39k*	93	2.20	1.95	0.493	1.60	0.568
	b: 40-49k	103	1.87	1.78		1.10	
	c: 50-69k	170	2.33	1.98		1.70	
	d: 70-99k	97	2.19	1.94		1.30	
	e: 100-149k	92	2.22	1.85		1.65	
	f: 150k+	44	1.93	1.93		1.30	
HHSize	a: 1	229	1.97	1.84	0.308	1.20	0.185
	b: 2	151	2.23	2.01		1.40	
	c: 3	191	2.21	1.87		1.70	
	d: 4 or more	89	2.37	1.89		1.80	
Driver Mileage Last Year	a: 0-5k*	40	2.37	2.11	0.586	1.25	0.616
	b: 6-9k*	41	2.18	1.90		1.80	
	c: 10-12k*	141	2.09	1.85		1.40	
	d: 13-15k*	48	2.56	2.00		2.05	
	e: 16-19k*	45	2.40	2.10		1.70	
	f: 20-23k*	104	2.00	1.80		1.30	
	g: 25k+	170	2.32	1.95		1.75	

Table A.16: Estimated Wiedemann 99 Acceleration at 80 kph (cc9) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Acceleration at 80kph [m/s^2] - cc9							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.60	1.40	0.091	0.10	0.163
	M	261	0.81	1.75		0.10	
Age	a: 20-24	71	0.56	1.19	0.230	0.10	0.481
	b: 25-29	209	0.53	1.20		0.10	
	c: 30-34	61	0.84	1.96		0.10	
	d: 35-39	80	0.96	2.00		0.20	
	e: 40-44	66	0.52	1.19		0.10	
	f: 45-59*	47	1.03	1.93		0.10	
	g: 60-69*	71	0.61	1.64		0.10	
	h: 70+*	56	0.84	1.75		0.10	
Race	Caucasian	453	0.77	1.67	0.033	0.10	0.084
	Not Caucasian	205	0.49	1.24		0.10	
Education	a: No college degree	191	0.70	1.59	0.965	0.10	0.363
	b: College degree	338	0.69	1.59		0.10	
	c: Graduate Degree	131	0.65	1.39		0.20	
Marital Status	a: single	219	0.57	1.27	0.349	0.10	0.736
	b: unmarried partners	49	0.47	1.18		0.10	
	c: married	282	0.81	1.74		0.10	
	d: divorced	89	0.66	1.64		0.10	
	e: widow(er)	19	0.94	1.92		0.20	
Income	a: Under 39k*	93	0.60	1.21	0.899	0.10	0.968
	b: 40-49k	103	0.66	1.55		0.10	
	c: 50-69k	170	0.59	1.31		0.10	
	d: 70-99k	97	0.80	1.74		0.10	
	e: 100-149k	92	0.69	1.69		0.10	
	f: 150k+	44	0.76	1.78		0.10	
HHSize	a: 1	229	0.64	1.54	0.774	0.10	0.813
	b: 2	151	0.79	1.70		0.10	
	c: 3	191	0.64	1.42		0.10	
	d: 4 or more	89	0.73	1.62		0.10	
Driver Mileage Last Year	a: 0-5k*	40	0.89	1.91	0.431	0.20	0.524
	b: 6-9k*	41	0.95	1.85		0.20	
	c: 10-12k*	141	0.78	1.76		0.10	
	d: 13-15k*	48	0.89	1.67		0.10	
	e: 16-19k*	45	0.78	1.62		0.20	
	f: 20-23k*	104	0.82	1.82		0.15	
	g: 25k+	170	0.49	1.19		0.10	

Table A.17: Estimated Wiedemann 99 Desired Velocity (v_des) Coefficient Segmented by Driver Attributes

Wiedemann 99 - Desired Travel Speed [m/s] - vdes							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	32.65	3.16	**0.002**	33.30	**0.000**
	M	261	31.90	2.72		31.50	
Age	a: 20-24	71	31.98	2.56	**0.000**	31.80	**0.000**
	b: 25-29	209	33.55	2.69		34.10	
	c: 30-34	61	33.42	2.07		33.40	
	d: 35-39	80	31.41	1.85		31.20	
	e: 40-44	66	34.67	1.87		35.35	
	f: 45-59*	47	29.79	2.86		29.60	
	g: 60-69*	71	30.67	2.84		31.10	
	h: 70+*	56	30.07	3.66		30.30	
Race	Caucasian	453	31.83	2.98	**0.000**	31.80	**0.000**
	Not Caucasian	205	33.51	2.76		34.10	
Education	a: No college degree	191	32.23	3.15	**0.000**	32.80	**0.000**
	b: College degree	338	31.86	2.79		31.80	
	c: Graduate Degree	131	33.79	2.96		34.90	
Marital Status	a: single	219	33.00	2.92	**0.000**	33.50	**0.000**
	b: unmarried partners	49	35.29	1.49		35.50	
	c: married	282	32.12	2.89		32.35	
	d: divorced	89	30.81	1.89		31.00	
	e: widow(er)	19	27.98	3.94		26.90	
Income	a: Under 39k*	93	31.40	3.75	**0.000**	31.80	**0.000**
	b: 40-49k	103	33.27	2.80		33.90	
	c: 50-69k	170	32.87	2.58		33.20	
	d: 70-99k	97	30.69	3.05		30.90	
	e: 100-149k	92	33.83	2.77		35.05	
	f: 150k+	44	32.36	2.79		32.40	
HHSize	a: 1	229	31.94	3.27	**0.000**	31.70	**0.000**
	b: 2	151	31.39	3.06		31.80	
	c: 3	191	33.02	2.34		33.20	
	d: 4 or more	89	33.61	2.86		34.90	
Driver Mileage Last Year	a: 0-5k*	40	31.36	2.25	**0.000**	31.55	**0.000**
	b: 6-9k*	41	29.05	3.43		29.00	
	c: 10-12k*	141	31.69	2.54		31.80	
	d: 13-15k*	48	31.90	3.70		32.45	
	e: 16-19k*	45	33.81	2.93		34.80	
	f: 20-23k*	104	30.59	2.41		30.90	
	g: 25k+	170	34.05	2.25		34.50	

Table A.18: Estimated Gipps Desired Acceleration (a_des) Coefficient Segmented by Driver Attributes

Gipps - Desired Acceleration [m/s^2] - a_des							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	1.42	0.90	**0.032**	1.30	**0.001**
	M	261	1.26	1.00		1.00	
Age	a: 20-24	71	1.42	0.74	**0.001**	1.40	**0.000**
	b: 25-29	209	1.44	0.90		1.30	
	c: 30-34	61	1.61	1.18		1.20	
	d: 35-39	80	1.37	1.03		1.00	
	e: 40-44	66	1.27	0.74		1.10	
	f: 45-59*	47	1.05	0.75		0.90	
	g: 60-69*	71	1.41	1.07		1.20	
	h: 70+*	56	0.93	0.95		0.60	
Race	Caucasian	453	1.19	0.90	**0.000**	1.00	**0.000**
	Not Caucasian	205	1.71	0.92		1.60	
Education	a: No college degree	191	1.09	0.84	**0.000**	0.90	**0.000**
	b: College degree	338	1.52	0.97		1.40	
	c: Graduate Degree	131	1.31	0.92		1.10	
Marital Status	a: single	219	1.63	0.95	**0.000**	1.50	**0.000**
	b: unmarried partners	49	1.28	0.68		1.10	
	c: married	282	1.20	0.93		1.00	
	d: divorced	89	1.39	0.97		1.10	
	e: widow(er)	19	0.62	0.55		0.40	
Income	a: Under 39k*	93	1.18	0.77	**0.000**	1.10	**0.000**
	b: 40-49k	103	1.79	1.08		1.70	
	c: 50-69k	170	1.19	0.84		1.00	
	d: 70-99k	97	1.51	0.98		1.40	
	e: 100-149k	92	1.32	0.92		1.10	
	f: 150k+	44	1.32	0.94		1.25	
HHSize	a: 1	229	1.48	1.01	**0.037**	1.30	0.068
	b: 2	151	1.36	1.07		1.10	
	c: 3	191	1.21	0.78		1.10	
	d: 4 or more	89	1.34	0.79		1.30	
Driver Mileage Last Year	a: 0-5k*	40	1.38	1.06	0.217	1.15	0.238
	b: 6-9k*	41	1.06	0.97		0.80	
	c: 10-12k*	141	1.28	0.99		1.00	
	d: 13-15k*	48	1.40	1.04		1.40	
	e: 16-19k*	45	1.33	1.02		1.00	
	f: 20-23k*	104	1.38	0.93		1.10	
	g: 25k+	170	1.15	0.73		1.00	

Table A.19: Estimated Gipps Desired Deceleration (d_des) Coefficient Segmented by Driver Attributes

Gipps - Desired Deceleration [m/s^2] - d_des							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	-2.54	0.91	**0.000**	-2.60	**0.000**
	M	261	-2.87	0.94		-3.10	
Age	a: 20-24	71	-2.62	0.79	**0.000**	-2.60	**0.000**
	b: 25-29	209	-2.70	0.86		-2.80	
	c: 30-34	61	-2.23	0.86		-2.20	
	d: 35-39	80	-2.92	0.86		-3.10	
	e: 40-44	66	-3.30	0.73		-3.50	
	f: 45-59*	47	-2.54	0.92		-2.50	
	g: 60-69*	71	-2.48	1.04		-2.60	
	h: 70+*	56	-2.39	1.16		-2.50	
Race	Caucasian	453	-2.72	0.92	**0.024**	-2.90	**0.025**
	Not Caucasian	205	-2.55	0.94		-2.60	
Education	a: No college degree	191	-2.77	0.92	**0.000**	-2.90	**0.000**
	b: College degree	338	-2.49	0.92		-2.50	
	c: Graduate Degree	131	-2.99	0.88		-3.20	
Marital Status	a: single	219	-2.49	0.88	**0.000**	-2.40	**0.000**
	b: unmarried partners	49	-3.46	0.54		-3.60	
	c: married	282	-2.71	0.88		-2.80	
	d: divorced	89	-2.73	1.02		-3.00	
	e: widow(er)	19	-1.85	1.21		-1.60	
Income	a: Under 39k*	93	-2.61	1.03	**0.000**	-2.90	**0.000**
	b: 40-49k	103	-2.32	0.87		-2.30	
	c: 50-69k	170	-2.78	0.81		-2.90	
	d: 70-99k	97	-2.48	1.04		-2.40	
	e: 100-149k	92	-3.09	0.83		-3.30	
	f: 150k+	44	-2.47	0.85		-2.40	
HHSize	a: 1	229	-2.54	1.01	**0.000**	-2.70	**0.000**
	b: 2	151	-2.53	0.93		-2.60	
	c: 3	191	-2.74	0.83		-2.80	
	d: 4 or more	89	-3.10	0.79		-3.30	
Driver Mileage Last Year	a: 0-5k*	40	-2.94	0.82	**0.000**	-2.80	**0.000**
	b: 6-9k*	41	-2.15	1.07		-2.00	
	c: 10-12k*	141	-2.45	0.96		-2.40	
	d: 13-15k*	48	-2.55	1.00		-2.65	
	e: 16-19k*	45	-2.80	0.92		-3.00	
	f: 20-23k*	104	-2.65	1.01		-2.90	
	g: 25k+	170	-3.02	0.77		-3.20	

Table A.20: Estimated Gipps Following Vehicle Perception of Leading Vehicle Desired Deceleration (d_lead) Coefficient Segmented by Driver Attributes

Gipps - Perception of Leading Vehicle Desired Deceleration [m/s^2] - d_lead							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	-2.44	0.93	**0.007**	-2.40	**0.003**
	M	261	-2.64	0.95		-2.80	
Age	a: 20-24	71	-2.51	0.88	**0.000**	-2.50	**0.000**
	b: 25-29	209	-2.61	0.90		-2.70	
	c: 30-34	61	-2.13	0.89		-2.00	
	d: 35-39	80	-2.79	0.85		-2.85	
	e: 40-44	66	-3.04	0.81		-3.20	
	f: 45-59*	47	-2.32	0.88		-2.20	
	g: 60-69*	71	-2.29	0.97		-2.20	
	h: 70+*	56	-2.10	1.10		-2.05	
Race	Caucasian	453	-2.60	0.95	**0.001**	-2.70	**0.001**
	Not Caucasian	205	-2.35	0.91		-2.30	
Education	a: No college degree	191	-2.66	0.95	**0.000**	-2.80	**0.000**
	b: College degree	338	-2.35	0.94		-2.30	
	c: Graduate Degree	131	-2.75	0.87		-2.90	
Marital Status	a: single	219	-2.32	0.90	**0.000**	-2.30	**0.000**
	b: unmarried partners	49	-3.09	0.67		-3.20	
	c: married	282	-2.59	0.92		-2.70	
	d: divorced	89	-2.60	1.01		-2.70	
	e: widow(er)	19	-1.75	1.09		-1.70	
Income	a: Under 39k*	93	-2.46	1.02	**0.000**	-2.40	**0.000**
	b: 40-49k	103	-2.14	0.89		-2.10	
	c: 50-69k	170	-2.77	0.88		-2.80	
	d: 70-99k	97	-2.21	0.90		-2.20	
	e: 100-149k	92	-2.78	0.91		-3.00	
	f: 150k+	44	-2.41	0.86		-2.50	
HHSize	a: 1	229	-2.35	0.97	**0.000**	-2.30	**0.000**
	b: 2	151	-2.35	0.95		-2.40	
	c: 3	191	-2.70	0.90		-2.80	
	d: 4 or more	89	-2.85	0.80		-3.00	
Driver Mileage Last Year	a: 0-5k*	40	-2.72	0.73	**0.000**	-2.80	**0.000**
	b: 6-9k*	41	-2.07	1.04		-1.90	
	c: 10-12k*	141	-2.30	0.98		-2.20	
	d: 13-15k*	48	-2.31	1.05		-2.30	
	e: 16-19k*	45	-2.62	0.93		-2.80	
	f: 20-23k*	104	-2.55	1.00		-2.70	
	g: 25k+	170	-2.89	0.82		-3.00	

Table A.21: Estimated Minimum Desired Gap at a Stop (g_min) Coefficient Segmented by Driver Attributes

Gipps - Minimum Gap at a Stop [m] - g_min							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	4.22	2.93	0.801	3.85	0.865
	M	261	4.29	3.31		3.60	
Age	a: 20-24	71	4.53	2.75	**0.000**	4.00	**0.000**
	b: 25-29	209	3.75	2.76		3.50	
	c: 30-34	61	4.20	3.08		3.20	
	d: 35-39	80	3.18	3.21		2.15	
	e: 40-44	66	4.77	3.34		4.25	
	f: 45-59*	47	4.51	3.05		4.60	
	g: 60-69*	71	5.16	3.11		5.00	
	h: 70+*	56	5.34	3.45		4.75	
Race	Caucasian	453	3.90	3.14	**0.000**	3.20	**0.000**
	Not Caucasian	205	4.98	2.83		4.80	
Education	a: No college degree	191	4.02	3.18	0.387	3.30	0.319
	b: College degree	338	4.30	2.95		4.00	
	c: Graduate Degree	131	4.48	3.26		4.00	
Marital Status	a: single	219	4.60	2.74	**0.000**	4.20	**0.000**
	b: unmarried partners	49	5.19	3.26		5.10	
	c: married	282	3.88	3.10		3.20	
	d: divorced	89	3.69	3.29		2.70	
	e: widow(er)	19	5.90	3.64		5.00	
Income	a: Under 39k*	93	4.55	3.13	**0.012**	3.90	**0.004**
	b: 40-49k	103	4.80	2.60		4.70	
	c: 50-69k	170	3.64	3.08		3.00	
	d: 70-99k	97	4.65	3.07		4.00	
	e: 100-149k	92	4.76	3.23		4.85	
	f: 150k+	44	4.11	3.20		3.10	
HHSize	a: 1	229	4.49	3.05	0.079	4.20	0.050
	b: 2	151	4.29	3.11		3.60	
	c: 3	191	3.79	3.05		3.10	
	d: 4 or more	89	4.58	3.12		4.30	
Driver Mileage Last Year	a: 0-5k*	40	3.61	3.00	0.468	3.00	0.373
	b: 6-9k*	41	4.73	3.31		4.00	
	c: 10-12k*	141	4.26	3.08		3.40	
	d: 13-15k*	48	4.44	3.09		3.50	
	e: 16-19k*	45	4.64	3.20		4.20	
	f: 20-23k*	104	3.93	3.27		3.25	
	g: 25k+	170	3.92	3.17		3.40	

Table A.22: Estimated Gipps Driver Reaction Time (t_rxn) Coefficient Segmented by Driver Attributes

Gipps - Reaction Time [s] - t_rxn							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.69	0.52	0.904	0.50	0.934
	M	261	0.70	0.52		0.50	
Age	a: 20-24	71	0.76	0.50	**0.000**	0.50	**0.000**
	b: 25-29	209	0.53	0.41		0.40	
	c: 30-34	61	0.97	0.57		1.00	
	d: 35-39	80	0.61	0.47		0.50	
	e: 40-44	66	0.62	0.47		0.50	
	f: 45-59*	47	0.90	0.61		0.80	
	g: 60-69*	71	0.75	0.52		0.60	
	h: 70+*	56	0.90	0.61		0.85	
Race	Caucasian	453	0.66	0.52	**0.006**	0.50	**0.002**
	Not Caucasian	205	0.77	0.51		0.80	
Education	a: No college degree	191	0.54	0.49	**0.000**	0.40	**0.000**
	b: College degree	338	0.77	0.52		0.80	
	c: Graduate Degree	131	0.73	0.49		0.70	
Marital Status	a: single	219	0.80	0.50	**0.000**	0.80	**0.000**
	b: unmarried partners	49	0.67	0.47		0.50	
	c: married	282	0.63	0.52		0.50	
	d: divorced	89	0.61	0.48		0.50	
	e: widow(er)	19	0.98	0.64		1.10	
Income	a: Under 39k*	93	0.81	0.53	**0.000**	0.80	**0.000**
	b: 40-49k	103	0.85	0.49		0.90	
	c: 50-69k	170	0.46	0.39		0.40	
	d: 70-99k	97	0.86	0.57		0.80	
	e: 100-149k	92	0.71	0.51		0.50	
	f: 150k+	44	0.92	0.60		0.90	
HHSize	a: 1	229	0.74	0.52	**0.000**	0.70	**0.000**
	b: 2	151	0.82	0.52		0.80	
	c: 3	191	0.53	0.47		0.40	
	d: 4 or more	89	0.72	0.50		0.70	
Driver Mileage Last Year	a: 0-5k*	40	0.68	0.48	**0.000**	0.50	**0.000**
	b: 6-9k*	41	0.83	0.55		0.80	
	c: 10-12k*	141	0.89	0.55		0.90	
	d: 13-15k*	48	0.83	0.62		0.80	
	e: 16-19k*	45	0.72	0.42		0.70	
	f: 20-23k*	104	0.62	0.49		0.50	
	g: 25k+	170	0.51	0.47		0.40	

Table A.23: Estimated Gipps Desired Velocity (v_des) Coefficient Segmented by Driver Attributes

Gipps - Desired Travel Speed [m/s] - v_des							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	32.95	3.31	**0.000**	33.40	**0.000**
	M	261	31.68	3.76		31.50	
Age	a: 20-24	71	31.94	2.64	**0.000**	31.60	**0.000**
	b: 25-29	209	33.84	3.05		34.10	
	c: 30-34	61	33.67	2.06		33.70	
	d: 35-39	80	31.18	3.37		31.30	
	e: 40-44	66	34.62	2.78		35.15	
	f: 45-59*	47	30.15	2.88		30.30	
	g: 60-69*	71	30.49	4.42		31.40	
	h: 70+*	56	30.21	3.64		30.85	
Race	Caucasian	453	32.03	3.80	**0.000**	31.90	**0.000**
	Not Caucasian	205	33.41	2.68		33.90	
Education	a: No college degree	191	32.46	4.19	**0.000**	33.00	**0.000**
	b: College degree	338	31.91	3.01		31.95	
	c: Graduate Degree	131	33.83	3.48		34.90	
Marital Status	a: single	219	33.01	2.95	**0.000**	33.50	**0.000**
	b: unmarried partners	49	34.95	2.77		35.40	
	c: married	282	32.32	4.03		32.70	
	d: divorced	89	31.00	1.73		31.20	
	e: widow(er)	19	28.11	4.07		28.20	
Income	a: Under 39k*	93	31.79	3.90	**0.000**	31.70	**0.000**
	b: 40-49k	103	33.04	2.83		33.70	
	c: 50-69k	170	33.26	3.87		33.40	
	d: 70-99k	97	30.57	3.69		30.90	
	e: 100-149k	92	33.69	3.29		34.90	
	f: 150k+	44	32.64	3.20		32.65	
HHSize	a: 1	229	32.06	3.18	**0.000**	31.90	**0.000**
	b: 2	151	31.37	4.02		31.90	
	c: 3	191	33.46	2.79		33.30	
	d: 4 or more	89	33.12	4.34		34.70	
Driver Mileage Last Year	a: 0-5k*	40	31.83	2.97	**0.000**	31.65	**0.000**
	b: 6-9k*	41	28.97	3.45		29.40	
	c: 10-12k*	141	32.01	2.83		32.00	
	d: 13-15k*	48	31.93	3.91		32.70	
	e: 16-19k*	45	33.83	3.01		34.30	
	f: 20-23k*	104	30.34	4.25		31.15	
	g: 25k+	170	34.35	2.83		34.90	

Table A.24: Estimated Intelligent Driver Model Maximum Desired Acceleration (a) Coefficient Segmented by Driver Attributes

Intelligent Driver Model - Maximum Desired Acceleraton [m/s^2] - a							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.98	0.73	**0.070**	0.8	**0.000**
	M	261	0.86	0.97		0.5	
Age	a: 20-24	71	1.08	0.59	0.095	1	**0.000**
	b: 25-29	209	0.95	0.72		0.7	
	c: 30-34	61	0.92	0.64		0.8	
	d: 35-39	80	1.02	1.06		0.6	
	e: 40-44	66	0.88	0.96		0.5	
	f: 45-59*	47	0.68	0.63		0.5	
	g: 60-69*	71	1.03	0.98		0.8	
	h: 70+*	56	0.73	1.02		0.3	
Race	Caucasian	453	0.78	0.80	**0.000**	0.5	**0.000**
	Not Caucasian	205	1.28	0.81		1.3	
Education	a: No college degree	191	0.66	0.71	**0.000**	0.4	**0.000**
	b: College degree	338	1.11	0.83		1	
	c: Graduate Degree	131	0.89	0.90		0.5	
Marital Status	a: single	219	1.20	0.70	**0.000**	1.2	**0.000**
	b: unmarried partners	49	0.93	0.99		0.5	
	c: married	282	0.76	0.80		0.5	
	d: divorced	89	0.98	0.99		0.6	
	e: widow(er)	19	0.32	0.32		0.2	
Income	a: Under 39k*	93	0.82	0.61	**0.000**	0.7	**0.000**
	b: 40-49k	103	1.39	0.87		1.4	
	c: 50-69k	170	0.67	0.57		0.5	
	d: 70-99k	97	1.11	0.92		0.9	
	e: 100-149k	92	0.97	1.01		0.5	
	f: 150k+	44	0.91	0.87		0.65	
HHSize	a: 1	229	1.08	0.86	**0.000**	0.9	**0.002**
	b: 2	151	0.94	0.90		0.6	
	c: 3	191	0.73	0.63		0.5	
	d: 4 or more	89	1.00	0.96		0.6	
Driver Mileage Last Year	a: 0-5k*	40	0.98	0.93	**0.006**	0.7	**0.001**
	b: 6-9k*	41	0.59	0.60		0.3	
	c: 10-12k*	141	0.91	0.79		0.7	
	d: 13-15k*	48	0.92	0.97		0.55	
	e: 16-19k*	45	0.97	0.84		0.7	
	f: 20-23k*	104	1.00	0.95		0.7	
	g: 25k+	170	0.68	0.69		0.5	

Table A.25: Estimated Intelligent Driver Model Maximum Desired Deceleration (b) Coefficient Segmented by Driver Attributes

Intelligent Driver Model - Maximum Desired Deceleraton [m/s^2] - b							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	2.07	1.43	**0.000**	1.8	**0.000**
	M	261	2.82	1.33		3.5	
Age	a: 20-24	71	1.78	1.24	**0.000**	1.4	**0.000**
	b: 25-29	209	2.24	1.44		2.2	
	c: 30-34	61	1.71	1.29		1.3	
	d: 35-39	80	3.01	1.32		4	
	e: 40-44	66	3.16	1.11		3.9	
	f: 45-59*	47	2.21	1.46		2	
	g: 60-69*	71	2.31	1.54		2.1	
	h: 70+*	56	2.60	1.45		2.9	
Race	Caucasian	453	2.66	1.36	**0.000**	3.1	**0.000**
	Not Caucasian	205	1.70	1.39		1.1	
Education	a: No college degree	191	2.82	1.27	**0.000**	3.2	**0.000**
	b: College degree	338	1.94	1.45		1.4	
	c: Graduate Degree	131	2.80	1.34		3.2	
Marital Status	a: single	219	1.62	1.31	**0.000**	1.1	**0.000**
	b: unmarried partners	49	3.18	1.07		3.9	
	c: married	282	2.65	1.35		3.05	
	d: divorced	89	2.82	1.43		3.6	
	e: widow(er)	19	2.36	1.59		2.5	
Income	a: Under 39k*	93	2.13	1.43	**0.000**	1.8	**0.000**
	b: 40-49k	103	1.41	1.23		0.9	
	c: 50-69k	170	2.60	1.32		2.9	
	d: 70-99k	97	2.16	1.54		2	
	e: 100-149k	92	2.85	1.34		3.5	
	f: 150k+	44	2.40	1.38		2.35	
HHSize	a: 1	229	2.14	1.55	**0.004**	1.8	**0.005**
	b: 2	151	2.31	1.44		2.3	
	c: 3	191	2.53	1.30		2.6	
	d: 4 or more	89	2.70	1.37		3.2	
Driver Mileage Last Year	a: 0-5k*	40	2.92	1.20	**0.000**	3.3	**0.000**
	b: 6-9k*	41	2.10	1.46		2	
	c: 10-12k*	141	2.09	1.39		1.7	
	d: 13-15k*	48	2.24	1.54		1.85	
	e: 16-19k*	45	2.35	1.40		2.4	
	f: 20-23k*	104	2.65	1.54		3.55	
	g: 25k+	170	2.96	1.16		3.4	

Table A.26: Estimated Intelligent Driver Model Free Acceleration (δ) Coefficient Segmented by Driver Attributes

Intelligent Driver Model - Free Acceleration Component [unitless] - delta							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	45.97	38.76	0.479	31	0.360
	M	261	43.80	38.06		27	
Age	a: 20-24	71	50.69	41.57	0.233	42	0.261
	b: 25-29	209	42.52	37.19		26	
	c: 30-34	61	47.64	38.15		29	
	d: 35-39	80	54.30	39.16		58	
	e: 40-44	66	40.79	33.33		28.5	
	f: 45-59*	47	39.72	37.47		25	
	g: 60-69*	71	43.97	41.65		27	
	h: 70+*	56	42.88	39.90		26	
Race	Caucasian	453	40.65	37.90	**0.000**	22	**0.000**
	Not Caucasian	205	54.42	37.99		50	
Education	a: No college degree	191	33.82	34.66	**0.000**	18	**0.000**
	b: College degree	338	51.21	39.67		44.5	
	c: Graduate Degree	131	46.11	37.17		36	
Marital Status	a: single	219	52.25	39.38	**0.000**	46	**0.000**
	b: unmarried partners	49	45.63	34.28		38	
	c: married	282	36.71	35.74		19.5	
	d: divorced	89	59.43	40.94		71	
	e: widow(er)	19	24.63	28.57		18	
Income	a: Under 39k*	93	42.48	38.67	**0.003**	22	**0.004**
	b: 40-49k	103	54.26	37.49		49	
	c: 50-69k	170	36.12	36.34		17.5	
	d: 70-99k	97	50.08	39.13		42	
	e: 100-149k	92	43.38	36.72		29	
	f: 150k+	44	43.07	39.43		25.5	
HHSize	a: 1	229	54.30	40.58	**0.000**	53	**0.004**
	b: 2	151	40.91	37.62		23	
	c: 3	191	38.61	36.25		21	
	d: 4 or more	89	42.94	34.95		30	
Driver Mileage Last Year	a: 0-5k*	40	34.85	35.56	**0.016**	20	**0.022**
	b: 6-9k*	41	35.83	35.95		24	
	c: 10-12k*	141	41.72	38.34		22	
	d: 13-15k*	48	54.83	40.75		59.5	
	e: 16-19k*	45	43.67	40.58		20	
	f: 20-23k*	104	52.05	39.99		41.5	
	g: 25k+	170	39.14	35.75		22.5	

Table A.27: Estimated Intelligent Driver Model Desired Time Gap (t_gap) Coefficient Segmented by Driver Attributes

Intelligent Driver Model - Desired Time Gap [s] - t_gap							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	0.73	0.57	**0.007**	0.6	**0.001**
	M	261	0.62	0.47		0.5	
Age	a: 20-24	71	0.75	0.40	**0.000**	0.7	**0.000**
	b: 25-29	209	0.49	0.20		0.5	
	c: 30-34	61	1.01	0.38		1	
	d: 35-39	80	0.54	0.42		0.4	
	e: 40-44	66	0.56	0.39		0.5	
	f: 45-59*	47	0.84	0.53		0.8	
	g: 60-69*	71	0.79	0.65		0.7	
	h: 70+*	56	1.06	1.12		0.75	
Race	Caucasian	453	0.67	0.56	0.475	0.6	0.103
	Not Caucasian	205	0.70	0.47		0.6	
Education	a: No college degree	191	0.65	0.67	**0.025**	0.5	**0.000**
	b: College degree	338	0.74	0.45		0.6	
	c: Graduate Degree	131	0.60	0.50		0.5	
Marital Status	a: single	219	0.68	0.39	**0.000**	0.6	**0.000**
	b: unmarried partners	49	0.52	0.39		0.5	
	c: married	282	0.67	0.45		0.6	
	d: divorced	89	0.62	0.56		0.4	
	e: widow(er)	19	1.73	1.52		1.5	
Income	a: Under 39k*	93	0.86	0.88	**0.000**	0.6	**0.000**
	b: 40-49k	103	0.77	0.43		0.7	
	c: 50-69k	170	0.54	0.25		0.5	
	d: 70-99k	97	0.76	0.59		0.7	
	e: 100-149k	92	0.64	0.49		0.5	
	f: 150k+	44	0.94	0.61		0.9	
HHSize	a: 1	229	0.69	0.70	**0.049**	0.5	**0.000**
	b: 2	151	0.77	0.46		0.8	
	c: 3	191	0.61	0.32		0.6	
	d: 4 or more	89	0.66	0.52		0.5	
Driver Mileage Last Year	a: 0-5k*	40	0.57	0.42	**0.000**	0.45	**0.000**
	b: 6-9k*	41	1.20	1.21		0.7	
	c: 10-12k*	141	0.90	0.48		0.9	
	d: 13-15k*	48	0.78	0.58		0.6	
	e: 16-19k*	45	0.47	0.35		0.4	
	f: 20-23k*	104	0.66	0.52		0.5	
	g: 25k+	170	0.51	0.26		0.5	

Table A.28: Estimated Intelligent Driver Model Jam Distance (g_min) Coefficient Segmented by Driver Attributes

Intelligent Driver Model - Jam Distance [m] - g_min							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	3.91	3.03	0.659	3.4	0.188
	M	261	3.80	3.44		2.4	
Age	a: 20-24	71	4.08	3.03	**0.025**	3.6	**0.049**
	b: 25-29	209	3.42	2.75		3	
	c: 30-34	61	4.53	3.34		3.6	
	d: 35-39	80	3.52	3.19		2.6	
	e: 40-44	66	4.66	3.57		4.3	
	f: 45-59*	47	4.64	3.49		4.5	
	g: 60-69*	71	3.48	3.34		2.6	
	h: 70+*	56	4.00	3.64		2.1	
Race	Caucasian	453	3.53	3.26	**0.000**	2.5	0.000
	Not Caucasian	205	4.56	2.92		4.2	
Education	a: No college degree	191	3.41	3.20	0.068	2.3	**0.033**
	b: College degree	338	4.07	3.11		3.65	
	c: Graduate Degree	131	4.00	3.38		2.8	
Marital Status	a: single	219	4.33	2.89	**0.000**	3.9	**0.000**
	b: unmarried partners	49	5.28	3.42		6.1	
	c: married	282	3.41	3.22		2.35	
	d: divorced	89	3.23	3.16		2.1	
	e: widow(er)	19	4.04	3.86		2.2	
Income	a: Under 39k*	93	4.10	3.37	**0.006**	3.3	**0.001**
	b: 40-49k	103	4.63	2.72		4.2	
	c: 50-69k	170	3.25	3.02		2.4	
	d: 70-99k	97	3.87	3.32		2.8	
	e: 100-149k	92	4.40	3.39		3.75	
	f: 150k+	44	3.38	3.45		1.65	
HHSize	a: 1	229	3.94	3.09	0.054	3.6	0.078
	b: 2	151	4.07	3.43		3.1	
	c: 3	191	3.37	3.03		2.6	
	d: 4 or more	89	4.37	3.33		3	
Driver Mileage Last Year	a: 0-5k*	40	3.34	3.13	0.792	2.55	0.718
	b: 6-9k*	41	3.20	3.40		2	
	c: 10-12k*	141	3.99	3.38		3.3	
	d: 13-15k*	48	3.67	3.28		2.5	
	e: 16-19k*	45	3.87	3.35		2.8	
	f: 20-23k*	104	3.75	3.32		2.8	
	g: 25k+	170	3.96	3.26		3	

Table A.29: Estimated Intelligent Driver Model Desired Velocity (v_des) Coefficient Segmented by Driver Attributes

Intelligent Driver Model - Desired Velocity [m/s] - v_des							
Attributes	Categories	N	Mean	Std. Dev	ANOVA p-value	Median	Kruskal Wallis p-value
Gender	F	400	33.70	3.68	**0.002**	34.1	**0.000**
	M	261	32.84	3.36		32.1	
Age	a: 20-24	71	33.05	3.06	**0.000**	32.5	**0.000**
	b: 25-29	209	34.68	3.59		34.9	
	c: 30-34	61	34.56	2.42		34.3	
	d: 35-39	80	32.02	2.35		31.5	
	e: 40-44	66	35.47	2.81		35.85	
	f: 45-59*	47	30.57	3.35		30.5	
	g: 60-69*	71	31.74	3.32		31.7	
	h: 70+*	56	31.32	3.95		31	
Race	Caucasian	453	32.93	3.73	**0.000**	32.5	**0.000**
	Not Caucasian	205	34.35	2.99		34.6	
Education	a: No college degree	191	33.54	3.68	**0.000**	33.9	**0.000**
	b: College degree	338	32.66	3.29		32.5	
	c: Graduate Degree	131	34.91	3.67		35.7	
Marital Status	a: single	219	33.81	3.53	**0.000**	34.2	**0.000**
	b: unmarried partners	49	35.78	2.86		36	
	c: married	282	33.43	3.65		33.4	
	d: divorced	89	31.51	2.11		31.4	
	e: widow(er)	19	29.37	3.86		29.7	
Income	a: Under 39k*	93	32.68	4.05	**0.000**	32.5	**0.000**
	b: 40-49k	103	33.65	3.67		34.3	
	c: 50-69k	170	34.33	3.26		34.3	
	d: 70-99k	97	31.63	3.41		31.4	
	e: 100-149k	92	34.95	3.40		35.7	
	f: 150k+	44	33.41	3.64		33.05	
HHSize	a: 1	229	32.87	3.48	**0.000**	32.3	**0.000**
	b: 2	151	32.22	3.90		32.4	
	c: 3	191	34.37	3.01		34.2	
	d: 4 or more	89	34.39	3.60		35.7	
Driver Mileage Last Year	a: 0-5k*	40	32.52	3.23	**0.000**	32.1	**0.000**
	b: 6-9k*	41	30.43	4.08		29.7	
	c: 10-12k*	141	32.71	3.02		32.8	
	d: 13-15k*	48	32.56	5.28		33.15	
	e: 16-19k*	45	34.91	3.38		35.3	
	f: 20-23k*	104	31.32	2.68		31.3	
	g: 25k+	170	35.26	2.92		35.7	

Bibliography

114th Congress. Fixing America's Surface Transportation (FAST) Act, Pub. L. No. 114-94 (2015). Retrieved from https://www.congress.gov/114/crpt/hrpt357/CRPT-114hrpt357.pdf

Abbas, M., Higgs, B., Adam, Z., & Medina, A. (2011). Comparison of car-following models when calibrated to individual drivers using naturalistic data. *Proceedings of the 90th Annual Meeting of the Transportation Research Board Compendium of Papers*.

Abbas, M., Higgs, B., & Medina, A. (2011). Naturalistic car driver behavior: A hybrid wiedemann-GHR model calibration. *Proceedings of the 90th Annual Meeting of the Transportation Research Board Compendium of Papers*.

Abbas, M., Higgs, B., Medina, A., & Yang, C. Y. D. (2010). A Hybrid Wiedemann-GHR Model Calibration using Naturalistic Driving Data. *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems*.

Abbas, M., Medina, A., Chong, L., Higgs, B., McGhee, C., Fontain, M., … Ova, K. (2012). *Driver behavior in traffic. Final Report: FHWA-HRT-12-036*. Blacksburg, VA.

Aghabayk, K., Sarvi, M., Forouzideh, N., & Young, W. (2013). New car-following model considering impacts of multiple lead vehicle types. *Transportation Research Record*, (2390), 131–137. https://doi.org/10.3141/2390-14

Aghabayk, K., Sarvi, M., & Young, W. (2014). Attribute selection for modelling driver's car-following behaviour in heterogeneous congested traffic conditions. *Transportmetrica A: Transport Science*, *10*(March 2015), 457–468. https://doi.org/10.1080/23249935.2013.787558

Ahmed, M. M. (2016). *Driver performance and behavior in adverse weather conditions: An investigation using the SHRP2 naturalistic driving study data*. *Proposal: SHRP2 Implementation Assistance Program - Round 4*. Laramie, WY.

Ahmed, M. M., Ghasemzadeh, A., Hammit, B., Khan, N., Das, A., Ali, E., … Eldeeb, H. (2018). *Driver performance and behavior in adverse weather conditions: An investigation using the SHRP2 naturalistic driving study data: Phase 2 final report*. *WY-18/05F*. Laramie, Wyoming.

Alexiadis, V., Jeannotte, K., & Chandra, A. (2004). *Traffic Analysis Toolbox Volume I: Traffic Analysis Tools Primer*. *Publication No. FHWA-HRT-04-038* (Vol. I).

Barbosa, L. de C. (1961). *Studies of traffic flow models*. *Reports No. 202A-1*. Columbus, OH.

Berthaume, A. L., James, R. M., Hammit, B. E., Foreman, C., & Melson, C. (2018). Variations in driver behavior: An analysis of car-following behavior heterogeneity

as a function of road type and traffic condition. *Transportation Research Record, 2672*(37), 31–44. https://doi.org/https://doi.org/10.1177/0361198118798713

Berthaume, Andrew L. (2015). *Microscopic modeling of driver behavior based on modifying field theory for work zone application*. University of Massachusetts - Amherst.

Bexelius, S. (1968). An extended model for car-following. *Transportation Research, 2*(1), 13–21.

Blatt, A., Pierowicz, J., Flanigan, M., Lin, P.-S., Kourtellis, A., Jovanis, P., … Hoover, M. (2015). *Naturalistic Driving Study: Field Data Collection*. *National Academcy of Sciences*. https://doi.org/10.17226/22367

Boyce, T. E., & Geller, E. S. (2002). An instrumented vehicle assessment of problem behavior and driving style: Do younger males really take more risks? *Accident Analysis and Prevention*, *34*(1), 51–64. https://doi.org/10.1016/S0001-4575(00)00102-0

Brackstone, M. (2003). Driver psychological types and car following: is there a correlation? results of a pilot study. *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, (1993), 245–250.

Brackstone, M., & McDonald, M. (1999). Car-following: A historical review. *Transportation Research Part F: Traffic Psychology and Behaviour*, *2*(4), 181–196.

Brackstone, M., Waterson, B., & McDonald, M. (2009). Determinants of following headway in congested traffic. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(2), 131–142. https://doi.org/10.1016/j.trf.2008.09.003

Brockfeld, E., Kuhne, R. D., & Wagner, P. (2004). Calibration and validation of microscopic traffic flow models. *Transportation Research Record: Journal of the Transportation Research Board*, *1876*, 62–70.

Brockfeld, E., & Wagner, P. (2006). Validating microscopic flow models. *Proceedings of the Institute of Electrical and Electronics Engineers Intelligent Transportation Systems Conference*, 1605–1608.

Campbell, K. L. (n.d.). The SHRP2 Naturalistic Driving Study: Addressing Driver Performance and Behavior in Traffic Safety. *Transportation Research News*, 30–36.

Chandler, R. E., Herman, R., & Montroll, E. W. (1958). Traffic Dynamics: Studies in Car Following. *Operations Research*, *6*(2), 165–184. https://doi.org/10.1287/opre.6.2.165

Chen, C., Li, L., Hu, J., & Geng, C. (2010). Calibration of MITSIM and IDM car-following model based on NGSIM trajectory datasets. *Proceedings of 2010 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2010*, 48–53. https://doi.org/10.1109/ICVES.2010.5550943

Chen, K., & Yu, L. (2007). Microscopic traffic-emission simulation and case study for evaluation of traffic control strategies. *Journal of Transportation Systems Engineering and Information Technology*, *7*(1), 93–99.

Chong, L., Abbas, M., Medina Flintsch, A., & Higgs, B. (2013). A rule-based neural network approach to model driver naturalistic behavior in traffic. *Transportation Research Part C: Emerging Technologies*, *32*, 207–223. https://doi.org/10.1016/j.trc.2012.09.011

Ciuffo, B., Punzo, V., & Montanino, M. (2012). *The calibration of traffic simulation models: Report on the assessment of different goodness of fit measures and optimization algorithms*. *JCR Scientific and Technical Reports*. Luxembourg. https://doi.org/10.2788/7975

Constantinescu, Z., Marinoiu, C., & Vladoiu, M. (2010). Driving Style Analysis Using Data Mining Techniques. *International Journal of Computers Communications & Control*, *5*(5), 654. https://doi.org/10.15837/ijccc.2010.5.2221

Corbett, C. (2007). Vehicle-related crime and the gender gap. *Psychology Crime & Law*, *13*(245–263).

de Waard, D., Dijksterhuis, C., & Brookhuis, K. A. (2009). Merging into heavy motorway traffic by young and elderly drivers. *Accident Analysis & Prevention*, *41*, 588–597.

Dijker, T., Bovy, P. H. L., & Vermijs, R. G. M. M. (1998). Car-following under congested conditions: Empirical findings. *Transportation Research Record: Journal of the Transportation Research Board*, *1644*(1), 20–28. https://doi.org/10.3141/1644-03

Dowling, R., Skabardonis, A., & Alexiadis, V. (2004). *Traffic Analysis Toolbox Volume III : Guidelines for Applying Traffic Microsimulation Modeling Software*. *Publication No. FHWA-HRT-04-040* (Vol. III).

Durrani, U., Lee, C., & Maoh, H. (2016). Calibrating the Wiedemann's vehicle-following model using mixed vehicle-pair interactions. *Transportation Research Part C: Emerging Technologies*, *67*, 227–242. https://doi.org/10.1016/j.trc.2016.02.012

Elander, J., West, R., & French, D. (1993). Behavioral Correlates of Individual Differences in Road-Traffic Crash Risk: An Examination of Methods and Findings. *Psychological Bulletin*, *113*(2), 279–294. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=paovfta&NEWS=N&AN=00006823-199303000-00004

Ericsson, E. (2000). Variability in urban driving patterns. *Transportation Research Part D: Transport and Environment*, *5*(5), 337–354. https://doi.org/10.1016/S1361-9209(00)00003-1

European Cooperation in Science and Technology. (n.d.). MULTITUDE: Methods and

tools for supporting the Use caLibration and vadaTIon of Traffic simUlation moDEls. Retrieved January 30, 2018, from www.multitude-project.eu

Evans, L., & Wasielewski, P. (1983). Risky driving related to driver and vehicle characteristics. *Accident Analysis and Prevention*, *15*(2), 121–136. https://doi.org/10.1016/0001-4575(83)90068-4

Fan, R., Wang, W., Liu, P., & Yu, H. (2013). Using VISSIM simulation model and Surrogate Safety Assessment Model for estimating field measured traffic conflicts at freeway merge areas. *IET Intelligent Transport Systems*, *7*(1), 68–77. https://doi.org/10.1049/iet-its.2011.0232

Federal Highway Administration. (2013). Performance Management. Retrieved July 10, 2017, from https://www.fhwa.dot.gov/map21/factsheets/pm.cfm

Federal Highway Administration Office of Operations. (2017). Next Generation Simulation (NGSIM). Retrieved November 11, 2017, from https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm

Federal Highway Administration Office of Operations Research and Development. (2017a). Moving FHWA Work Zone Driver Model towards Practical Application (Ongoing). Retrieved from https://www.fhwa.dot.gov/research/tfhrc/projects/operations/ams/project.cfm#03c

Federal Highway Administration Office of Operations Research and Development. (2017b). Narrowing Freeway Lanes and Shoulders to Create Additional Travel Lanes (Ongoing). Retrieved August 1, 2017, from https://www.fhwa.dot.gov/research/tfhrc/projects/operations/ams/project.cfm

Federal Highway Administration Office of Planning. (2007). *The Transportation Planning Process Key Issues: A Briefing Book for Transportation Decisionmakers, Officials, and Staff. Publication No. FHWA-HEP-07-039*. Washington, D.C. https://doi.org/10.1017/CBO9781107415324.004

Fernandez, S., & Ito, T. (2016). Driver classification for intelligent transportation systems using fuzzy logic. *Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 1212–1216. https://doi.org/10.1109/ITSC.2016.7795711

Fritzsche, H. T. (1994). A model for traffic simulation. *Traffic Engineering and Control*, *35*(5), 317–321.

Gazis, D. C., Herman, R., & Rothery, R. W. R. (1961). Nonlinear follow-the-leader models of traffic flow. *Operations Research*, *9*(4), 545–567. https://doi.org/10.1287/opre.9.4.545

Geng, X., Liang, H., Xu, H., & Yu, B. (2016). Influences of leading-vehicle types and environmental conditions on car-following behavior. In *International Federation of Automatic Control* (Vol. 49, pp. 151–156). Elsevier B.V.

https://doi.org/10.1016/j.ifacol.2016.07.724

Gettman, D., Pu, L., Sayed, T., & Shelby, S. (2008). *Surrogate Safety Assessment Model and Validation: Final Report*. *Publication No. FHWA-HRT-08-051*.

Ghasemzadeh, A., Hammit, B. E., Ahmed, M. M., & Eldeeb, H. (2018). Complementary methodologies to identify weather conditions in naturalistic driving study trips : Lessons learned from the SHRP2 Naturalistic Driving Study & Roadway Information Database. *Proceedings of the 97th Annual Meeting of the Transportation Research Board Compendium of Papers*.

Gipps, P. G. (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, *15*(2), 105–111.

Gorman, T., Stowe, L., & Hankey, J. (2015). *S31: NDS data dissemination activities, Task 1.6: radar post-processing*.

Habtemichael, F. G., & De Picado Santos, L. (2014). Crash risk evaluation of aggressive driving on motorways: Microscopic traffic simulation approach. *Transportation Research Part F: Traffic Psychology and Behaviour*, *23*, 101–112. https://doi.org/10.1016/j.trf.2013.12.022

Halkias, J., & Colyar, J. (2006). *Next Generation SIMulation (NGSIM) Fact Sheet*. *FHWA-HRT-06-135*. Washington DC. Retrieved from https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm

Hammit, B. E., Ghasemzadeh, A., James, R. M., Ahmed, M. M., & Young, R. K. (2018). Evaluation of weather-related freeway car-following behavior using the SHRP2 Naturalistic Driving Study database. *Transportation Research Part F*, *59*, 244–259. https://doi.org/doi.org/10.1016/j.trf.2018.08.023

Hammit, B. E., James, R. M., & Ahmed, M. (2018a). A case for online traffic simulation: Systematic procedure to calibrate car-following models using vehicle data. *21st Institute of Electrical and Electronics Engineers International Conference on Intelligent Transportation Systems (ITSC)*, 3785–3790. https://doi.org/10.1109/ITSC.2018.8569684

Hammit, B. E., James, R. M., & Ahmed, M. (2018b). Radar-vision algorithms to process the trajectory-level driving data in the SHRP2 Naturalistic Driving Study. *21st Institute of Electrical and Electronics Engineers International Conference on Intelligent Transportation Systems (ITSC)*, 2126–2133. https://doi.org/10.1109/ITSC.2018.8569688

Hammit, B. E., James, R. M., Ahmed, M., & Young, R. (2019). Towards the development of weather-dependent microsimulation models. *Transportation Research Record*. https://doi.org/https://doi.org/10.1177/0361198119844743

Hankey, J. M., McClafferty, J. A., & Perez, M. A. (2016). *Description of the SHRP2 Naturalistic Database and the Crash, Near Crash, and Baseline Data Sets*. *The*

Second Strategic Highway Research Program, Transportation Research Board of the National Academies. Retrieved from https://vtechworks.lib.vt.edu/bitstream/handle/10919/70850/SHRP_2_CrashNearCrashBaselineReport_4-25-16.pdf?sequence=1

Hao, H., Ma, W., & Xu, H. (2016). A fuzzy logic-based multi-agent car-following model. *Transportation Research Part C: Emerging Technologies*, *69*, 477–496. https://doi.org/10.1016/j.trc.2015.09.014

Higgs, B., & Abbas, M. (2013). A two-step segmentation algorithm for behavioral clustering of naturalistic driving styles. *IEEE 16th International Conference on Intelligent Transportation Systems (ITSC)*, 857–862.

Higgs, B., & Abbas, M. (2014). Multi-resolution comparison of car-following models using naturalistic data. *Proceedings of the 93rd Annual Meeting of the Transportation Research Board Compendium of Papers*.

Higgs, B., & Abbas, M. (2015). Segmentation and clustering of car-following behavior: Recognition of driving patterns. *IEEE Transactions on Intelligent Transportation Systems*, *16*(1), 81–90. https://doi.org/10.1109/TITS.2014.2326082

Higgs, B., Abbas, M., & Medina, A. (2011). Analysis of the wiedemann car-following model over different speeds using naturalistic data. In *Road Safety and Simulation* (pp. 1–22). Indianapolis, Indiana. Retrieved from http://onlinepubs.trb.org/onlinepubs/conferences/2011/RSS/3/Higgs,B.pdf

Higgs, B., Abbas, M., & Medina, A. (2012). Naturalistic car driver behavior : A comparison of car-following models. *Proceedings of the 19th ITS World Congress Conference Proceedings*, 1–10.

Hoogendoorn, S. P., & Hoogendoorn, R. G. (2010). Generic calibration framework for joint estimation of car-following models by using microscopic data. *Transportation Research Record: Journal of the Transportation Research Board*, *2188*, 37–45. https://doi.org/10.3141/2188-05

Hoogendoorn, S. P., & Ossen, S. (2006). Empirical analysis of two-leader car-following behavior. *European Journal of Transport and Infrastucture Reseach (EJTIR)*, *6*(3), 229–246.

Hoogendoorn, S. P., Ossen, S., & Schreuder, M. (2006). Empirics of multianticipative car-following behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *1965*, 112–120. https://doi.org/10.3141/1965-12

Hoogendoorn, S. P., Ossen, S., & Schreuder, M. (2007). Multi-anticipative car-following behavior: An emperical analysis. In A. Schadschneider, T. Poschel, R. Kuhne, M. Schreckenberg, & E. Wolf, Dietrich (Eds.), *Traffic and Granular Flow '05* (pp. 687–697). Washington D.C.

Insurance Information Institute. (2018). Facts + Statistics: Highway safety. Retrieved

July 10, 2018, from https://www.iii.org/fact-statistic/facts-statistics-highway-safety

Ishibashi, M., Okuwa, M., Doi, S., & Akamatsu, M. (2007). Indices for characterizing driving style and their relevance to car following behavior. *Proceedings of the SICE Annual Conference*, 1132–1137. https://doi.org/10.1109/SICE.2007.4421155

Itkonen, T. H., Pekkanen, J., Lappi, O., Kosonen, I., Luttinen, T., & Summala, H. (2017). Trade-off between jerk and time headway as an indicator of driving style. *PLoS ONE*, *12*(10), 1–20. https://doi.org/10.1371/journal.pone.0185856

James, R. M., & Hammit, B. E. (2019). Identifying contributory factors to heterogeneity in driving behavior : A clustering and classification approach. *Transportation Research Record*.

James, R. M., Hammit, B. E., & Ahmed, M. (2018). Exploring the use of driver attributes to characterize heterogeneity in naturalistic driving behavior. *21st Institute of Electrical and Electronics Engineers International Conference on Intelligent Transportation Systems (ITSC)*, 1047–1052. https://doi.org/10.1109/ITSC.2018.8569497

James, R. M., Hammit, B. E., & Boyles, S. D. (2019). Methods to obtain representative car-following model parameters from trajectory-level data for use in microsimulation. *Transportation Research Record*.

James, R. M., Melson, C., Hu, J., & Bared, J. (2019). Characterizing the impact of production adaptive cruise control on traffic flow: an investigation. *Transportmetrica B: Transport Dynamics*, *7*(1), 992–1012. https://doi.org/https://doi.org/10.1080/21680566.2018.1540951

Jeannotte, K., Chandra, A., Alexiadis, V., & Skabardonis, A. (2004). Traffic Analysis Toolbox Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools. *Publication No. FHWA-HRT-04-039*, *II*(July), 1–108.

Kesting, A., & Treiber, M. (2008). Calibrating car-following models by using trajectory data: Methodological study. *Transportation Research Record: Journal of the Transportation Research Board*, *2088*, 148–156. https://doi.org/10.3141/2088-16

Kesting, A., Treiber, M., & Helbing, D. (2010). Enhanced Intelligent Driver Model to Access the Impact of Driving Strategies on Traffic Capacity. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *368*(1928), 4585–4605. https://doi.org/10.1098/rsta.2010.0084

Khodayari, A., Ghaffari, A., Kazemi, R., & Braunstingl, R. (2012). A Modified Car-Following Model Based on a Neural Network Model of the Human Driver Effects. *IEEE Transactions on Systems Man and Cybernetics Part A: Systems and Humans*, *42*(6), 1440–1449. https://doi.org/Doi 10.1109/Tsmca.2012.2192262

Kim, J., & Mahmassani, H. S. (2011). Correlated parameters in driving behavior models.

Transportation Research Record: Journal of the Transportation Research Board, *2249*, 62–77. https://doi.org/10.3141/2249-09

Kleisen, L. M. B. (2011). *The relationship between thinking and driving styles and their contribution to young driver road safety*. University of Canberra, Bruce, Australia.

Kometani, E., & Sasaki, T. (1959). Dynamic behaviour of traffic with a nonlinear spacing-speed relationship. *Symposium on Theory of Traffic Flow*, 105–119.

Kurtc, V., & Treiber, M. (2016). Calibrating the Local and Platoon Dynamics of Car-following Models on the Reconstructed NGSIM Data. In *Traffic and Granular Flow'15* (pp. 515–522). Springer. https://doi.org/10.1007/978-3-319-33482-0_65

Laapotti, S., Keskinen, E., & Rajalin, S. (2003). Comparison of young male and female drivers' attitude and self-reported traffic behaviour in Finland in 1978 and 2001. *Journal of Safety Research*, *34*, 579–587.

Li, C., Dixit, V. V., & Hamdar, S. H. (2016). Estimating Risk Attitude and Risk Perception in Car Following. *Proceedings of the 95th Annual Meeting of the Transportation Research Board Compendium of Papers*, 1–18.

Li, G., Li, S. E., Cheng, B., & Green, P. (2017). Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities. *Transportation Research Part C: Emerging Technologies*, *74*, 113–125. https://doi.org/10.1016/j.trc.2016.11.011

Li, L., & Chen, X. (2017). Vehicle headway modeling and its inferences in macroscopic/microscopic traffic flow theory: A survey. *Transportation Research Part C: Emerging Technologies*, *76*, 170–188. https://doi.org/10.1016/j.trc.2017.01.007

Liu, G. (2016). W99 Car Following Model - How It Works. Retrieved November 1, 2017, from http://w99demo.com/

Lochrane, T. W. P. (2014). *A new multidimensional psycho-physical framework for modeling car-following in a freeway work zone*. University of Central Florida.

Lochrane, T. W. P., Al-Deek, H., Dailey, D. J., & Bared, J. (2014). Using a Living Laboratory to Support Transportation Research for a Freeway Work Zone. *Journal of Transportation Engineering*, *140*(7). https://doi.org/10.1061/(ASCE)TE.1943-5436.0000674

Lochrane, T. W. P., Al-Deek, H., Dailey, D. J., & Krause, C. (2015). Modeling Driver Behavior in Work and Nonwork Zones: Multidimensional Psychophysical Car-Following Framework. *Transportation Research Record*, (2490), 116–126. https://doi.org/10.3141/2490-13

Lochrane, T. W. P., Al-deek, H., Jiang, X., Highway, N., Dailey, D. J., & Shurbutt, J. (2017). Capturing Driver Behavior by Measuring Time Gap in Freeway Work Zones. *Proceedings of the 96th Annual Meeting of the Transportation Research*

Board Compendium of Papers.

Ma, J., Zhou, F., Huang, Z., & James, R. M. (2018). Hardware-in-the-loop testing of connected and automated vehicle applications: a use case for queue-aware signalized intersection approach and departure. *Transportation Research Record*, 0361198118793001.

Madi, M. Y. (2016). Investigating and calibrating the dynamics of vehicles in traffic micro-simulations models. *Proceedings of the 6th Transport Research Arena*, *14*, 1782–1791. https://doi.org/10.1016/j.trpro.2016.05.144

Manstetten, D., Krautter, W., & Schwab, T. (1997). Traffic simulation supporting urban control system development. *Presented at 4th World Congress on Intelligent Transport System, Berlin, Germany*.

McLaughlin, S., Hankey, J., & Dingus, T. (2009). Driver Measurement: Methods and Applications. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics: 8th International Conference, EPCE 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings* (pp. 404–413). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02728-4_43

Mehmood, A., & Easa, S. (2009). Modeling reaction time in car-following behaviour based on human factors. *... Journal of Applied Science, Engineering and ...*, *3*(9), 325–333. Retrieved from http://www.waset.org/publications/2937

Michaels, R. M. (1963). Perceptual factors in car following. *Proceedings of the Second International Symposium on the Theory of Traffic Flow*, 44–59.

Montanino, M., Ciuffo, B., & Punzo, V. (2012). Calibration of microscopic traffic flow models against time-series data. *Proceedings of the 15th International Institute of Electrical and Electronics Engineers Conference on Intelligent Transportation Systems (ITSC)*, 108–114.

Montanino, M., & Punzo, V. (2013). Making NGSIM Data Usable for Studies on Traffic Flow Theory. *Transportation Research Record: Journal of the Transportation Research Board*, *2390*, 99–111. https://doi.org/10.3141/2390-11

Montanino, M., & Punzo, V. (2015). Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns. *Transportation Research Part B: Methodological*, *80*, 82–106. https://doi.org/10.1016/j.trb.2015.06.010

Monteil, J., Billot, R., Sau, J., Buisson, C., & El Faouzi, N.-E. (2014). Calibration, Estimation, and Sampling Issues of Car-Following Parameters. *Transportation Research Record: Journal of the Transportation Research Board*, *2422*, 131–140. https://doi.org/10.3141/2422-15

Netherlands Organisation for Scientific Research. (n.d.). Tracing congestion dynamics: with innovative microscopic data to better theory. Retrieved July 10, 2017, from

https://www.nwo.nl/en/research-and-results/research-projects/i/70/570.html

Olstam, J. J., & Tapani, A. (2004). Comparison of car-following models. *Swedish National Road and Transportation Research Institute, No. VTI re*(1), 116–127. https://doi.org/10.3141/1678-15

Ossen, S. (2008). *Longitudinal driving behavior: Theory and empirics*. Technische Universiteit Delft.

Ossen, S., & Hoogendoorn, S. P. (2005). Car-following behavior analysis from microscopic trajectory data. *Transportation Research Record: Journal of the Transportation Research Board*, *1934*, 13–21. https://doi.org/10.3141/1934-02

Ossen, S., & Hoogendoorn, S. P. (2007). Driver heterogeneity in car-following and its impact on modeling traffic dynamics. *Transportation Research Record: Journal of the Transportation Research Board*, *1999*, 95–103. https://doi.org/10.3141/1999-11

Ossen, S., & Hoogendoorn, S. P. (2008). Validity of trajectory-based calibration approach of car-following models in presence of measurement errors. *Transportation Research Record: Journal of the Transportation Research Board*, *2088*, 117–125. https://doi.org/10.3141/2088-13

Ossen, S., & Hoogendoorn, S. P. (2009). Reliability of parameter values estimated using trajectory observations. *Transportation Research Record: Journal of the Transportation Research Board*, *2124*, 36–44. https://doi.org/10.3141/2124-04

Ossen, S., & Hoogendoorn, S. P. (2011). Heterogeneity in car-following behavior: Theory and empirics. *Transportation Research Part C: Emerging Technologies*, *19*(2), 182–195. https://doi.org/10.1016/j.trc.2010.05.006

Ossen, S., Hoogendoorn, S. P., & Gorte, B. G. H. (2006). Interdriver differences in car-following: A vehicle trajectory-based study. *Transportation Research Record: Journal of the Transportation Research Board*, (1965), 121–129.

Panwai, S., & Dia, H. (2005a). A reactive agent-based neural network car following model. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2005*, 326–331. https://doi.org/10.1109/ITSC.2005.1520069

Panwai, S., & Dia, H. (2005b). Comparative evaluation of microscopic car-following behavior. *IEEE Transactions on Intelligent Transportation Systems*, *6*(3), 314–325. https://doi.org/10.1109/TITS.2005.853705

Papathanasopoulou, V., & Antoniou, C. (2015). Towards data-driven car-following models. *Transportation Research Part C: Emerging Technologies*, *55*, 496–509. https://doi.org/10.1016/j.trc.2015.02.016

Papathanasopoulou, V., & Antoniou, C. (2016). Flexible car following models incorporating information from adjacent lanes. *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*.

Papathanasopoulou, V., Markou, I., & Antoniou, C. (2016). Online calibration for microscopic traffic simulation and dynamic multi-step prediction of traffic speed. *Transportation Research Part C: Emerging Technologies*, *68*, 144–159. https://doi.org/10.1016/j.trc.2016.04.006

Punzo, V., Borzacchiello, M. T., & Ciuffo, B. (2011). On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data. *Transportation Research Part C: Emerging Technologies*, *19*(6), 1243–1262. https://doi.org/10.1016/j.trc.2010.12.007

Punzo, V., Ciuffo, B., & Montanino, M. (2012). Can results of car-following model calibration based on trajectory data be trusted? *Transportation Research Record: Journal of the Transportation Research Board*, *2315*, 89–99. https://doi.org/10.3141/2315-02

Punzo, V., & Montanino, M. (2016). Speed or spacing? Cumulative variables, and convolution of model errors and time in traffic flow models validation and calibration. *Transportation Research Part B: Methodological*, *91*, 21–33.

Punzo, V., Montanino, M., & Ciuffo, B. (2015). Do we really need to calibrate all the parameters? Variance-based sensitivity analysis to simplify microscopic traffic flow models. *IEEE Transactions on Intelligent Transportation Systems*, *16*(1), 184–193. https://doi.org/10.1109/TITS.2014.2331453

Punzo, V., & Simonelli, F. (2005). Analysis and comparison of microscopic traffic flow models with real traffic microscopic data. *Transportation Research Record: Journal of the Transportation Research Board*, *1934*, 53–63. https://doi.org/10.3141/1934-06

Ranjitkar, P., Nakatsuji, T., & Asano, M. (2004). Performance evaluation of microscopic traffic flow models with test track data. *Transportation Research Record: Journal of the Transportation Research Board*, *1876*, 90–100. https://doi.org/10.3141/1876-10

Ranney, T. A. (1999). Psychological factors that influence car-following and car-following model development. *Transportation Research Part F: Traffic Psychology and Behaviour*, *2*(4), 213–219. https://doi.org/10.1016/S1369-8478(00)00010-3

Roy, M. M., & Liersch, M. J. (2013). I am a better driver than you think: examining self-enhancement for driving ability. *Journal of Applied Social Psychology*, *43*(8), 1648–1659.

Saifuzzaman, M., & Zheng, Z. (2014). Incorporating human-factors in car-following models: A review of recent developments and research needs. *Transportation Research Part C: Emerging Technologies*, *48*, 379–403. https://doi.org/10.1016/j.trc.2014.09.008

Sangster, J., Rakha, H., & Du, J. (2013). Application of naturalistic driving data to modeling of driver car-following Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2390*, 20–33.

https://doi.org/10.3141/2390-03

Smadi, O. (2015). SHRP2 - Roadway information database. Retrieved April 30, 2018, from http://www.ctre.iastate.edu/shrp2-rid/

Song, G., Yu, L., & Zhang, Y. (2012). Applicability of traffic microsimulation models in vehicle emissions estimates: Case study of VISSIM. *Transportation Research Record: Journal of the Transportation Research Board*, *2270*, 132–141.

Soria, I., Elefteriadou, L., & Kondyli, A. (2014). Assessment of car-following models by driver type and under different traffic, weather conditions using data from an instrumented vehicle. *Simulation Modelling Practice and Theory*, *40*, 208–220.

Soria, I. S. (2010). *Assessment of car-following models using field data*. University of Florida.

Soria, I. S., & Elefteriadou, L. (2011). Assessment of car-following models using field data. *Proceedings of the 90th Annual Meeting of the Transportation Research Board Compendium of Papers*, 1–18.

Stevanovic, A., Stevanovic, J., Zhang, K., & Batterman, S. (2009). Optimizing traffic control to reduce fuel consumption and vehicular emissions: Integrated approach with VISSIM, CMEM, and VISGAOST. *Transportation Research Record: Journal of the Transportation Research Board*, *2128*, 105–113.

Tefft, B. C. (2017). Rates of Motor Vehicle Crashes, Injuries and Deaths in Relation to Driver Age, United States, 2014-2015. Retrieved July 1, 2018, from http://aaafoundation.org/rates-motor-vehicle-crashes-injuries-deaths-relation-driver-age-united-states-2014-2015/

Todosoiev, E. P. (1963). *The action point model of the driver-vehicle system*. *Report No. 202A-3*. Columbus, OH.

Toledo, T. (2007). Driving behaviour: Models and challenges. *Transport Reviews*, *27*(1), 65–84. https://doi.org/10.1080/01441640600823940

Transportation Research Board. (2010). *Highway Capacity Manual*. Washington D.C.

Treiber, M., Arne, K., & Helbing, D. (2006). Delays, inaccuracies, and anticipation in microscopic traffic models. *Physica A: Statistical Mechanics and Its Applications*, *360*(1), 71–88.

Treiber, M., Hennecke, A., & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, *62*(2), 1805–1824. https://doi.org/10.1103/PhysRevE.62.1805

Treiber, M., & Kesting, A. (2013a). Chapter 16: Calibration & Validation. In *Traffic Flow Dynamics* (pp. 303–337). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-32460-4_16

Treiber, M., & Kesting, A. (2013b). Microscopic calibration and validation of car-

following models – A systematic approach. *Procedia - Social and Behavioral Sciences*, *80*, 922–939. https://doi.org/10.1016/j.sbspro.2013.05.050

Ulleberg, P., & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. *Safety Science*, *41*(5), 427–443.

Underwood, G. (2013). On-road behaviour of younger and older novices during the first six months of driving. *Accident Analysis & Prevention*, *58*, 235–243.

US Department of Transportation Federal Highway Administration. (2018a). Concept to Countermeasure – Research to Deployment Using the SHRP2 Safety Data. Retrieved October 22, 2018, from https://www.fhwa.dot.gov/goshrp2/Solutions/All/NDS/Concept_to_Countermeasure__Research_to_Deployment_Using_the_SHRP2_Safety_Data

US Department of Transportation Federal Highway Administration. (2018b). Implementation Assistance Program. Retrieved October 22, 2018, from https://www.fhwa.dot.gov/goshrp2/ImplementationAssistance#round4

US Department of Transportation Federal Highway Administration. (2018c). Implementation Assistance Program.

US Department of Transportation Federal Highway Administration. (2018d). SHRP2 Solutions: Tools for the Road Ahead.

Virginia Tech Transportation Institute. (2018a). InSight: Background Information - Frequently Asked Questions. Retrieved from https://insight.shrp2nds.us/projectBackground/index

Virginia Tech Transportation Institute. (2018b). InSight SHRP2 NDS. Retrieved from https://insight.shrp2nds.us/

Virginia Tech Transportation Institute. (2018c). SHRP2 Naturalistic Driving Study Dataverse. Retrieved October 22, 2018, from https://insight.shrp2nds.us/dataverse/index

Wang, H., Wang, W., Chen, J., & Jing, M. (2011). Using trajectory data to analyze intradriver heterogeneity in car-following. *Transportation Research Record: Journal of the Transportation Research Board*, *2188*, 85–95. https://doi.org/10.3141/2188-10

Warshawsky-Livne, L., & Shinar, D. (2002). Effects of uncertainty, transmission type, driver age and gender on brake reaction and movement time. *Journal of Safety Research*, *33*(1), 117–128. https://doi.org/10.1016/S0022-4375(02)00006-3

Weng, J., Meng, Q., & Fwa, T. F. (2014). Vehicle headway distribution in work zones. *Transportmetrica A: Transport Science*, *10*(4), 285–303. https://doi.org/10.1080/23249935.2012.762564

Wiedemann, R. (1996). Original code of W99.

Wiedemann, R., & Reiter, U. (1992). *Microscopic traffic simulation: The simulation system MISSION, background and actual state*. *Project ICARUS (V1052) Final Report*.

Wilson, R. E. (2001). An analysis of gipps' car-following model of highway traffic. *Institute of Mathematics and Its Applications Journal of Applied Mathematics*, *66*, 509–537. Retrieved from http://hdl.handle.net/1983/506

Witten, I., Frank, E., Hall, M. A., & Pal, C. J. (2017a). Chapter 5: Credibility: Evaluating what's been learned. In *Data Mining: Practical Machine Learning Tools and Techniques* (Fourth Edi, pp. 169–208). Cambridge, MA: Elsevier.

Witten, I., Frank, E., Hall, M. A., & Pal, C. J. (2017b). Chapter 6: Trees and rules. In *Data Mining: Practical Machine Learning Tools and Techniques* (pp. 209–242).

Witten, I., Frank, E., Hall, M. A., & Pal, C. J. (2017c). Chapter 8: Data transformations. In *Data Mining: Practical Machine Learning Tools and Techniques* (pp. 285–334).

Witten, I., Frank, E., Hall, M. A., & Pal, C. J. (2017d). Chapter 9: Probabilistic methods. In *Data Mining: Practical Machine Learning Tools and Techniques* (pp. 335–416).

Wu, J., Brackstone, M., & McDonald, M. (2003). The validation of a microscopic simulation model: A methodological case study. *Transportation Research Part C: Emerging Technologies*, *11*(6), 463–479. https://doi.org/10.1016/j.trc.2003.05.001

Wu, J., Du, Y., Qi, G., & Xu, M. (2015). Leveraging longitudinal driving behaviour data with data mining techniques for driving style analysis. *IET Intelligent Transport Systems*, *9*(8), 792–801. https://doi.org/10.1049/iet-its.2014.0139

Xyntarakis, M., Alexiadis, V., Campbell, R., & Flanigan, E. (2016). *Active Transportation and Demand Management (ATDM) Trajectory-Level Validation State of the Practice Review*. *FHWA-JPO-14-193*.

Yang, C. Y. D., & Morton, T. (2012). *Trends of transportation simulation and modeling based on a selection of exploratory advanced research projects*. *USDOT Federal Highway Administration Exploratory Advanced Research Program*. Washington DC.

Ye, F., & Zhang, Y. (2009). Vehicle Type-Specific Headway Analysis Using Freeway Traffic Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2124*(1), 222–230. https://doi.org/10.3141/2124-22

Yin, S., Li, Z., Zhang, Y., Yao, D., Su, Y., & Li, L. (2009). Headway distribution modeling with regard to traffic status. *Proceedings of the IEEE Intelligent Vehicles Symposium*, 1057–1062. https://doi.org/10.1109/IVS.2009.5164427

Zheng, J., Suzuki, K., & Fujita, M. (2012). Evaluation of car-following models using trajectory data from real traffic. *Proceedings of the 8th International Conference on*

Traffic and Transportation Studies, 43, 356–366. https://doi.org/10.1016/j.sbspro.2012.04.109

Zhu, M., Wang, X., & Wang, X. (2016). Car-following headways in different driving situations: A naturalistic driving study. *Proceedings of the 16th COTA International Conference on Transportation Professionals*, 1419–1428.